2006

THE UTILITY OF THE STRUCTURED INVENTORY OF MALINGERED SYMPTOMATOLOGY AS A SCREEN FOR THE FEIGNING OF NEUROCOGNITIVE DEFICIT AND PSYCHOPATHOLOGY IN A CIVIL FORENSIC SAMPLE

Yvonne Renee Alwes
University of Kentucky, yrstah0@email.uky.edu

ABSTRACT OF THESIS

THE UTILITY OF
THE STRUCTURED INVENTORY OF MALINGERED SYMPTOMATOLOGY
AS A SCREEN FOR THE FEIGNING OF
NEUROCOGNITIVE DEFICIT AND PSYCHOPATHOLOGY
IN A CIVIL FORENSIC SAMPLE

Detection of malingering is a significant concern in forensic psychological assessments. The best-validated tests currently available are time-intensive for both test-takers and mental health professionals. Thus, well-validated, brief screening measures for malingering would be useful in a forensic environment. The Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997) has demonstrated potential in this role. The present study attempts replication of previous studies while extending validation from analogue and male criminal forensic samples to both men and women in a civil forensic setting. The SIMS' accuracy in the detection of both neurocognitive and psychiatric symptom feigning is evaluated by comparing its performance to stringent multi-scale criterion measures in a large forensic sample. Cut scores suggested by previous studies yield high sensitivity and negative predictive power in this sample when the SIMS is used to detect psychiatric symptom malingering; however, these cut scores perform inadequately here when screening for the feigning of neurocognitive impairment, and no alternative cut score functions well in this capacity. The results lend support to the utility of the SIMS as a screen for psychiatric symptom malingering by men and women in a civil forensic setting.

KEYWORDS: Malingering, Structured Inventory of Malingered Symptomatology, Forensic Assessment, Neurocognitive Feigning, Psychiatric Feigning

Yvonne Renee Alwes
May 18, 2006

THE UTILITY OF
THE STRUCTURED INVENTORY OF MALINGERED SYMPTOMATOLOGY
AS A SCREEN FOR THE FEIGNING OF
NEUROCOGNITIVE DEFICIT AND PSYCHOPATHOLOGY
IN A CIVIL FORENSIC SAMPLE

By

Yvonne Renee Alwes

David T. R. Berry, Ph. D.
Director of Thesis

David T. R. Berry, Ph. D.
Director of Graduate Studies

THESIS

Yvonne Renee Alwes

The Graduate School

University of Kentucky

2006

THE UTILITY OF
THE STRUCTURED INVENTORY OF MALINGERED SYMPTOMATOLOGY
AS A SCREEN FOR THE FEIGNING OF
NEUROCOGNITIVE DEFICIT AND PSYCHOPATHOLOGY
IN A CIVIL FORENSIC SAMPLE

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
College of Arts and Sciences
at the University of Kentucky

By

Yvonne Renee Alwes

Lexington, Kentucky

Director: Dr. David T. R. Berry, Professor of Psychology

Lexington, Kentucky

2006

To Don, the light and love of my life,
without whose loving encouragement and support this project would never have been *undertaken*

ACKNOWLEDGMENTS

The opportunity for me to conduct this investigation and the quality of this research project and resultant thesis are owed in no small measure to the support, efforts, and expert guidance of my mentor and Thesis Committee Chair, Dr. David T. R. Berry. My thanks go also to Dr. Robert P. Granacher, Jr., not only for providing access to the data which fueled this research, but for his helpful mentoring as well. Several fellow students contributed substantial time and effort assisting with data collection, coding, and entry. I am grateful to my in-laws, the Alwes Family, for their continual encouragement during this effort. I am, further, deeply indebted to Otto for his unwavering support, for assiduously sharing his inextinguishable hope – for his steadfast belief in me. And, ultimately, everything worthwhile that I have accomplished has been inspired by my husband Don's character and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FILES

Chapter One

Introduction

Malingering is defined as a V-Code (V65.2) in the DSM-IV-TR (APA, 2000, p. 739) as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs." Studies of malingering have yielded widely varying estimates of its prevalence, but available figures may be low due to the failure to detect successful malingerers. In forensic psychological assessment, feigning of neurocognitive deficits has been reported to range from 15% to 48% (Inman et al., 1998), while rates of malingering of psychiatric symptoms have been suggested to range from 20% to 45% (Lewis, Simcox, & Berry, 2002).

When malingerers go undetected, the damages to society are significant. Health, disability, and unemployment insurance fraud as well as awards from litigation result in costs to insurance companies and government that are passed on to employers, employees, tax-payers, and the self-insured. In the criminal arena, the failure to hold true offenders criminally responsible exposes the general public to greater risk. In both criminal and civil settings, malingerers diminish the credibility of legitimate claimants and defendants. Thus, for many important reasons, it is vital to address the possibility of malingered symptoms in all psychological evaluations conducted in a compensation-seeking context.

Forensic psychologists and psychiatrists are thus in need of accurate yet easily (and inexpensively) administered instruments for identification of malingering in forensic populations (Rogers, 1997). The Structured Inventory of Malingered Symptomatology (SIMS) is a relatively new, brief test designed to address this need. In the present study, the SIMS' usefulness in identifying both psychiatric and neurocognitive symptom feigning in a civil forensic population is evaluated.

Embedded Validity Scales

MMPI-2. Initial systematic efforts at detecting dissimulation in test response involved the use of validity scales within measures designed primarily to detect neurocognitive or psychiatric problems. The Minnesota Multiphasic Personality Inventory – 2nd Edition (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) is still the most widely used psychological instrument in forensic assessment, and several of its validity scales are well-researched and relatively well-supported for the detection of dissimulation of psychopathology. The MMPI-2 Infrequency scales (F – Infrequency, Fb – Back Infrequency, and F[p] – Infrequency-Psychopathology) assess endorsement of severe psychopathology and atypical symptoms or attitudes, and have shown moderate success in the detection of malingering. F and Fb items were rarely endorsed by the normative sample, and F(p) items were rarely endorsed even by people with psychiatric disorders. The VRIN (Variable Response Inconsistency) and TRIN (True Response Inconsistency) scales, both measures of consistency of responding, are used to rule out random responding as an explanation for elevated scores on Infrequency scales. The F-K index (subtraction of the raw K – Defensiveness scale score from the raw F scale score) has also been found useful in identifying people who report fabricated symptoms. However, difficulties in establishing cut scores on these scales that could be reliably used in individual assessments to determine malingering, as well as ambiguity with respect to which scale performs best, persist (Lewis et al., 2002).

WAIS-R and WAIS-III. In the neuropsychological testing domain, a similar pattern in the development of methods to detect symptom feigning is evident. Initially, existing tests used to assess neurocognitive functioning were examined in an attempt to identify scales useful in discriminating feigners from honest reporters. The Digit Span subtests of the Wechsler Adult Intelligence Scale – Revised (Wechsler, 1981) and, more recently, of the Wechsler Adult Intelligence Scale – Third Edition (Wechsler, 1997) have demonstrated utility in this area. However, high false positive rates and the lack of establishment of a consistent cut score remain problems for use of the Digit Span scale for the detection of neurocognitive malingering (Inman & Berry, 2000).

Dedicated Malingering Measures

SIRS. The obvious advantage to using scales within existing tests that are part of standard batteries for assessing psychopathology or neurocognitive problems is the absence of a need to add tests to the battery. However, the limited success experienced with embedded scales on existing tests led to the development of instruments specifically designed to detect invalid response sets such as malingering. The best-validated dedicated test of malingering of psychopathology to date is the Structured Interview of Reported Symptoms (SIRS; Rogers, 1997; Rogers, Bagby, & Dickens, 1992; Lewis et al., 2002). The SIRS consists of 172 true/false items organized into eight primary scales and five supplementary scales. The various primary scales assess endorsement frequency, atypical symptom presentation, and the extent to which reported symptoms are observed during the interview (Heinze & Purisch, 2001). The SIRS is a structured interview; unlike the more common self-report measures, it must be administered by a trained professional, requiring anywhere from 30 to 60 minutes (Rogers et al., 1992).

Neurocognitive Tests of Effort. Many neurocognitive tests designed specifically to detect malingering approach the task with a *floor effect* strategy (Heinze & Purisch, 2001). This involves presenting items with very low difficulty that are judged by the malingerer to be more difficult for individuals with neurocognitive impairment, but which in actuality are not sensitive to impairment. A common format for such tests is presentation of a stimulus (a series of numbers or letters, or a line drawing) for a few seconds, followed by removal of the stimulus for a few seconds, then re-presentation of the stimulus along with at least one incorrect foil. The test-taker is instructed to identify the stimulus that was presented initially by itself.

Three tests operating on this principle are the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Thompson, 1997), and the Letter Memory Test (LMT; Inman et al., 1998). Each of these is a computer-administered forced-choice recognition test designed to detect poor effort.

For all three of these tests, the proportion of correctly answered items is used to determine malingering. Due to the low difficulty level of the items, even patients with severe brain injuries average close to 100% correct; therefore, cut scores are generally in the 90% - 95% range. For these tests, a performance below the cut score indicates malingering.

SIMS. The Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997) was developed to serve as a brief screening instrument for the detection of potential malingering. It is a self-report test consisting of 75 true/false items. One strength of the SIMS is the variety of symptomatology covered; some items are intended to be sensitive to neurocognitive feigning while others focus on feigning of psychiatric symptoms. Thus, the SIMS consists of five scales: Psychosis (P), Affective Symptoms (Af), Neurologic Impairment (N), Amnestic Symptoms (Am), and Low Intelligence (LI). The P, Af, N, and Am scales assess atypical symptom presentation, while the LI scale consists of simple items expected to be failed only by those feigning intellectual deficit (Heinze & Purisch, 2001). In keeping with its intended function as a screen, the SIMS has initially demonstrated high negative predictive power (NPP; the proportion of individuals classified by the test as Honest who truly are not malingering). Used in this manner, failure of the SIMS is followed by administration of another measure with high positive predictive power (PPP; the proportion of individuals classified by the test as Malingering who truly are malingering), such as the SIRS (Lewis et al., 2002). This two-stage assessment process can take advantage of the strengths of different tests to attain improved classification accuracy (Meehl, 1955), increasing the already high PPP of the 2nd-stage measure by way of the increased base rate among 1st-stage "failers", while also resulting in an overall more efficient use of time and effort. In this example, the SIMS self-report measure that takes 10 to 15 minutes to administer would be given to everyone undergoing forensic assessment; the SIRS structured interview that requires 30 to 60 minutes of professional administration time would be given only to those who "fail" the SIMS.

The SIMS has been validated in three studies with undergraduates (Smith, 1992; Smith & Burger, 1997; Edens, Otto, & Dwyer, 1999), two studies with male criminal defendants claiming incompetence to stand trial or diminished criminal responsibility (Heinze & Purisch, 2001; Lewis et al., 2002), and one study with adolescent offenders (Rogers, Hinds, & Sewell, 1996). Understanding the limits of these studies requires familiarity with research designs most applicable to this area of study. Three common methods of investigating malingering are the simulation design, known-groups comparison, and differential prevalence design. Each approach affords different strengths and limitations.

Malingering Research Designs

Simulation/Analogue. Simulation studies employ an analogue of malingering by including at least one group of participants instructed to feign psychological symptoms. A control group consists of participants instructed to respond honestly to testing; the simulators and control participants may be sampled randomly from the same population. However, in the strongest simulation design, a clinical comparison group consisting of participants with genuine disorders is also included. The simulation design provides a controlled experimental manipulation; however, the generalizability of the analogue participants' behavior to that of malingerers in real-life situations is unknown. One basis for questioning this generalizability is the recognition that the potentially life-altering consequences of success or failure at malingering in the real world are hardly comparable to those for simulators in experimental studies.

Known-Groups Comparisons. Known-groups comparisons make use of existing methods for identifying malingerers to form a group of participants designated as malingering. Individuals in a clinical or applied setting are classified as malingering or honest using a gold standard criterion measure. Responses of these malingering participants are compared to those of a group of honest responders. The behavior of both honest and malingering participants in these real-life settings is considered generalizable to that of other people in comparable settings. The accuracy of the gold standard, however, imposes a ceiling on the validity of the experimental measure in question. Further, the absence of experimental manipulation of the independent variable precludes inference of a clear causal relationship between malingering attempts and test performances.

Differential Prevalence. Differential prevalence designs avoid the problem of accurate classification of participants into known groups. Participants are sampled from two populations which are expected to have different base rates of malingering. Participants are not assessed to determine their group membership. However, responses of participants in the two resulting groups are compared, and differences are presumed to relate to differences in base rates of malingering. Although this method shares the generalizability of known-groups comparisons, its lack of information regarding which individuals are malingering and which are honest means that classification accuracy statistics cannot be determined.

Simulation studies, therefore, present superior internal validity and questionable external validity. Known-groups comparisons yield high external validity, but at the expense of internal validity. Differential prevalence designs allow for such poor internal validity as to be of much less value in investigating malingering.

Previous SIMS Studies

Smith (1992). Smith (1992) developed the SIMS and conducted initial validation using eight groups of college student participants – each of six groups was instructed to attempt to simulate a condition corresponding to one of the six SIMS scales; one group was instructed to feign general impairment; and one group was instructed to respond honestly. The SIMS Total score demonstrated a .95 hit rate (overall classification accuracy) with a cut score of >16. Cut scores for individual scales were also established, but the Total score proved to be superior to all of the individual scales at distinguishing malingerers in all groups from honest responders. The lack of a clinical control group is a serious limitation of this study, however.

Rogers, Hinds, & Sewell (1996). Rogers, Hinds, & Sewell (1996) studied a sample of 53 adolescent offenders by way of a within-subjects analogue design. The authors suggested that the fifth-grade reading level of the SIMS is especially valuable for use in an adolescent population. All participants had received dual diagnoses and had been referred by the courts to a state hospital for psychological treatment. Each participant completed testing twice – once under standard instructions and once under instructions to feign one of three disorders. All the adolescents were paid to participate and were offered additional modest incentives for successful feigning. For some participants, the researcher was aware of the experimental condition during data collection.

The Total cut score of >16 established by Smith (1992) reportedly yielded .87 PPP and .62 NPP in this sample. The Total cut score determined to be optimal for this study's sample was >40; this cut score was reported to produce .94 NPP and only .49 PPP. Optimal scale cut scores reported for this sample were also much higher than those established by Smith (1992), and also yielded high NPP on the individual scale scores. These figures may well be in error, however, as raising the cut scores substantially should *increase* PPP and *decrease* NPP. Although in the Discussion section the authors asserted that high PPP is ideal for a screening measure, rather than high NPP, they did not explicate in the Results section either their concept of "optimal" (i.e., what results optimal cut scores would yield) or the method used to determine the optimal cut scores. Sensitivity and specificity figures were not provided in the article. The use of a sample of offenders referred for treatment may add to external validity, although the monetary incentives offered imply that the results of this testing did not influence their real-life outcomes. Rogers et al. (1996) stated that they had found evidence for the SIMS as a "promising screen" (p. 254). Due to the afore-mentioned issues, however, it is difficult to evaluate the results provided.

Smith & Burger (1997). Smith & Burger (1997) administered the SIMS to 476 college students who were instructed to either respond honestly or simulate symptomatology. Feigning participants were asked to simulate one of six specific types of conditions (corresponding to the six SIMS scales) or general impairment. In this effort to validate the SIMS, half the sample was used to determine optimal cut scores for each scale and for the Total, and those cut scores were applied to the other half of the sample. Simulators scored higher on all SIMS scales than honest responders. NPP and PPP were not reported, but the SIMS Total score yielded a hit rate of .95, sensitivity of .96, and specificity of .88 using the cut score established with the first half of the sample. Sensitivity and specificity for the individual scales ranged from .75 to .88 and from .52 to .91, respectively; hit rates ranged from .74 to .89.

Participants were provided with vignettes and asked to imagine themselves as criminal defendants facing the possibility of a "serious sentence" (p. 186). They were also cautioned to avoid detection, but the analogue to real-world situations which give rise to malingering is weak. It is unclear how "optimal" was defined by the authors when they selected the cut scores, which were not reported. In the Discussion section, the authors recommended using a Total cut score of >14. The base rate of malingering in this study was .88, which is quite high. The use of half the sample for developing cut scores and the other half for validating them was superior to merely choosing the score that was optimal for the entire sample. The experimental simulation design with random assignment to groups provided high internal validity. However, no clinical control group was included, and no screening of participants for pathology was performed.

Edens, Otto, & Dwyer (1999). Edens, Otto, & Dwyer (1999) conducted a simulation study with 196 university students. Each participant completed both the SIMS and the Symptom Checklist 90–Revised (SCL-90-R; Derogatis, 1977) two times – once under normal instructions and once under instructions to feign specific symptomatology. In the feigning condition, participants were advised to avoid detection as feigners and were promised awards to the most successful malingerers. The SIMS scales were found to be generally sensitive to malingering, but not specifically sensitive to particular types of malingered symptomatology. The authors suggested that overlap in the scales and the lack of specificity of symptoms feigned by malingerers contributed to this finding.

The authors used the cut scores recommended by Smith & Burger (1997), which included a Total score >14. (As mentioned previously, individual scale cut scores were not reported by Smith & Burger, so it is not clear where Edens et al. obtained their individual scale cut scores.) PPP and NPP were not reported. Sensitivity and specificity for the Total score were .96 and .91, respectively. When "any scale elevation" was used as an indicator of malingering, sensitivity increased to 1.00, but specificity sank to .51. The Total score, with a hit rate of .96, was found to discriminate between malingering and honest responders better than any individual scale and better than the "any elevated scale" criterion. When subscales were used individually to screen for malingering, they were only somewhat focused in their sensitivity and specificity on the symptomatologies each was designed to cover. When screening for general malingering, sensitivity of individual scales ranged from .68 to .89; specificity figures ranged from .71 to .92. The SIMS produced more false positives with participants with actual symptomatology; among individuals with a psychiatric history or SCL-90-R Global Severity Index score of at least 45, the Total score yielded sensitivity of .97 to 1.00 and specificity of .78 to .91.

The sample size was larger than those in some studies; this allowed for the formation of several small groups of participants. Each group was instructed to feign a specific type of symptomatology. Although this study lacks generalizability due to its within-subjects simulation design and the absence of a clinical control group, it lends important support to the use of the SIMS Total score as a screen for malingering.

Heinze & Purisch (2001). Heinze & Purisch (2001) tested 57 male inmates of a maximum security state hospital who were charged with felonies and undergoing evaluation for competency to stand trial. All were suspected of feigning due to "compelling clinical evidence" (p. 28). In all cases, a multi-disciplinary research team and the evaluating psychiatrist suspected the participants of malingering based on "clinical indicators and behavioral observations" (p. 29). Due to the presumed 100% base rate of feigning in this sample, the study focused on evaluation of the sensitivity of the SIMS and several other instruments. The SIMS Total score demonstrated a sensitivity of .87 using a cut score of >13. It was not clear how the cut score was chosen.

The sample was rather small, although it did consist entirely of clinical malingerers. The method of determining the initial likelihood of malingering (the criterion for inclusion in the study) is questionable. The staff who made these determinations may not be representative of mental health professional in general. The purpose of the study was not focused on thorough evaluation of the utility of the various instruments in the detection of malingering – much of the discussion revolved around identifying different types of malingering. The design of the study did not allow for the calculation of many useful statistics such as hit rate, NPP, and PPP, and so did not lend itself to a good overall assessment of the SIMS.

Lewis, Simcox, & Berry (2002). Lewis, Simcox, & Berry (2002) tested 55 male defendants undergoing pre-trial psychological evaluations to determine competency to stand trial or mental state at the time of the offense. Each participant was classified as honest, malingering, or indeterminate on the basis of SIRS results. Participants who did not meet criteria for either malingering or honest responding (those whose classifications were indeterminate) were not included in analyses of the performance of the SIMS. Feigners scored significantly higher on all SIMS scales than did honest responders. Using a cut score of >16, the SIMS Total score demonstrated 1.00 NPP and .54 PPP. Sensitivity was 1.00, while specificity was .61, and overall hit rate = .73. Statistics for the individual scales were not reported.

As in all known-groups comparisons, here the validity of the SIRS caps the potential validity of the SIMS as a detector of malingering. The sample was rather small and exclusively male; all participants were facing federal charges. A strength of this study was the use of a sample of actual criminal defendants whose psychological evaluations carried real-life relevance.

Summary. Table 1.1 summarizes the research published to date on the SIMS. Notable limitations of individual studies include researchers not blind to a subject's condition (Rogers et al., 1996), confusing or erroneous statistics (Rogers et al., 1996), and weak criterion measures (Heinze & Purisch, 2001). Of the three studies conducted with offender participants, one included only adolescents (Rogers et al., 1996), and another employed only a single subject group (Heinze & Purisch, 2001).

Weaknesses common to several studies include: lack of patient or offender participants in simulation studies; small sample sizes in studies including offenders/patients; no female offender/patient participants; unreported cut scores (especially for sub-scales); unreported classification accuracy statistics (especially PPP and NPP); and failure to indicate how cut scores were chosen. These last three deficiencies, in particular, create difficulties for researchers attempting to establish reliable cut scores.

From a review of these six studies, support for a specific Total cut score is not evident. Reported cut scores ranged from >13 to >40, and produced hit rates from .73 to .96; sensitivity from .87 to 1.00; specificity from .61 to .91; PPP from .49 to .87; and NPP from .62 to 1.00. Of particular interest, cut scores of >14 and >16 have yielded maximal NPP (1.00). These results suggest that the objective for the SIMS to serve as a screen by minimizing false negatives could be realized, if a reliable cut score could be established.

Present Study

As shown in Table 1.1, thus far no studies have been published evaluating the validity of the SIMS in a civil forensic sample, and no forensic samples have included female participants. Given the prevalence estimates of malingering and the associated costs, it is important for a test of malingering to demonstrate utility in a civil forensic setting. Further, its accuracy with women in a forensic setting must be empirically established. The present study examined the accuracy of the SIMS in predicting malingering in a sample of male and female plaintiffs in workers' compensation and personal injury lawsuits. Unlike in previous investigations, here the SIMS' ability to detect both neurocognitive and psychiatric feigning was assessed against criterion measures appropriate for each type of feigning.

Malingering and Honest groups were also examined for differences on demographic variables such as gender, age, race, marital status, and education level, which might confound group differences on the SIMS. Differences in types of symptoms claimed (pain, psychiatric, brain damage, physical, and medical) were assessed as well.

Table 1.1

Previous SIMS Studies

Author(s)	Design*	Subjects**	N	Cut >	HR†	Sens†	Spec†	PPP†	NPP†
Smith (1992)	S	undergrad	●‡	16	0.95	●‡	●‡	●‡	●‡
Rogers et al. (1996)	S w-s	adol off	53	16	●‡	●‡	●‡	0.87	0.62
				40	●‡	●‡	●‡	0.49	0.94
Smith & Burger (1997)	S	undergrad	476	14	0.95	0.96	0.88	●‡	●‡
Edens et al. (1999)	S w-s	undergrad	196	14	0.96	0.96	0.91	●‡	●‡
Heinze & Purisch (2001)	KG	max secur	57	13	●‡	0.87	●‡	●‡	●‡
Lewis et al. (2002)	KGC	crim def	55	16	0.73	1.00	0.61	0.54	1.00

*S = Simulation, w-s = within-subjects, KG = Known-Group, KGC = Known-Groups Comparison

**undergrad = undergraduate students, adol off = adolescent offenders, max secur = maximum security inmates, crim def = criminal defendants

†HR = Hit Rate, Sens = Sensitivity, Spec = Specificity, PPP = Positive Predictive Power, NPP = Negative Predictive Power

‡Figures not published

Chapter Two

Methods

Participants

Participants were 308 individuals who had filed workers' compensation or personal injury lawsuits and were assessed in a private forensic psychiatric practice in Lexington, Kentucky. Types of cases included: cases involving head injury (33.8%, n = 104); personal injury cases not involving head injury (9.7%, n = 30); workers' compensation cases not involving head injury (55.2%, n = 170); and fitness for duty evaluations (0.6%, n = 2). Although residence information was not available, most of the individuals assessed at the psychiatric practice were residents of Kentucky, and most others resided in neighboring states. Men (n = 203) comprised 65.9% of the sample. The average age was 41.31 years, and ages ranged from 18 to 71 (SD = 11.245). The sample was predominantly Caucasian (93.2%, n = 287); 6.5% (n = 20) were African-American, and 0.3 % (n = 1) belonged to an "Other Single Race". Educational level ranged from 0 to 25 years, with an average of 12.07 years (SD = 2.596). Marital statuses broke down as follows: 65.9% married (n = 203); 17.2% divorced (n = 53); 11.7% single (n = 36); 1.9% separated (n = 6); and 1.9% widowed (n = 6). Socioeconomic status information on the participants was not available. The psychiatric practice was hired by the plaintiff in 39.3% (n = 121) of the cases, and by the defendant in the remaining cases. All participants were seeking compensation.

Measures

A large number of assessment instruments were administered to each participant by the staff of the psychiatric practice as per their protocol. Results from the following measures are examined in this study:

Background Questionnaire. This 22-page form, designed by the staff of the forensic psychiatric practice where participants were examined, is completed by every individual assessed at the practice. Information collected via this form includes demographics as well as personal, medical, and legal history.

Minnesota Multiphasic Personality Inventory – 2^{nd} Edition (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989*).* The MMPI-2 is a 567-item true/false self-report test designed to assess the presence of various types of psychopathology. Its validity scales have been extensively tested in forensic environments, and the Infrequency scales – F, Fb, and F(p), along with F-K – have demonstrated validity in discriminating between honest and malingering test-takers (Berry, Baer, & Harris, 1991; Rogers, Sewell, & Salekin, 1994; Arbisi & Ben-Porath, 1998), although specific scales and cut scores have not demonstrated consistent utility across studies (Lewis et al., 2002).

Structured Interview of Reported Symptoms (SIRS; Rogers et al., 1992*).* The SIRS is a structured interview for assessing malingering. It consists of 172 true/false items that comprise eight primary scales (coefficient $\alpha = .86$) and five supplementary scales (coefficient $\alpha = .75$). The eight primary scales are: Blatant Symptoms (BL), Subtle Symptoms (SU), and Selectivity of Symptoms (SEL) scales, which assess endorsement frequency; Rare Symptoms (RS), Improbable and Absurd Symptoms (IA), Symptom Combinations (SC), and Severity of Symptoms (SEV) scales, which assess atypical symptom presentation; and the Reported versus Observed Symptoms (RO) scale, which measures the extent to which reported symptoms are observed during the interview (Heinze & Purisch, 2001). Performance on each primary scale is categorized as Honest, Indeterminate, Probable Feigning, or Definite Feigning. There are several ways to achieve the classification of overall feigning: (1) at least one primary scale score is in the Definite Feigning range; (2) at least three primary scale scores fall in the Probable Feigning range; or (3) the Total raw score is greater than 76. In order to be classified as responding honestly, none of the malingering criteria may be met, and at least six primary scales must fall in the Honest range. Results which fail to meet criteria for either malingering or honest responding are classified as "indeterminate". Classification procedures were designed by the test's authors to minimize false positive classifications. High overall classification accuracy and PPP greater than .95 have been found in relatively high base rate settings such as those in which malingering is generally assessed (Rogers, 1997; Rogers, Bagby, & Dickens, 1992). The SIRS has been validated with known malingerers and against MMPI-2 Infrequency scales.

Victoria Symptom Validity Test (VSVT; Slick et al., 1997*).* The VSVT is a 48-item, computer-administered, two-alternative, forced-choice test of recognition memory that is insensitive to neurocognitive impairment but sensitive to attempts to feign such deficits. The stimuli are series of numbers. Six VSVT scores are produced: Total Items Correct, Easy Items Correct, Difficult Items Correct, Easy Items Response Latency, Difficult Items Response Latency, and Right-Left Preference. The proportion of items answered correctly is computed and scores below the cut score indicate malingering. The VSVT has been validated in compensation-seeking as well as non-compensation-seeking samples for the detection of poor effort (Thompson, 2002). Coefficient α = .82 to .89; test-retest reliability = .53 to .84; sensitivity (proportion of malingerers successfully detected) = .83; and specificity (proportion of non-malingerers successfully classified) = .95.

Test of Memory Malingering (TOMM; Tombaugh, 1996*)*. The TOMM is another two-alternative, forced-choice visual recognition test developed to test inadequate effort by test-takers. It is computer-administered, and consists of 50 items composed of line-drawing stimuli. The TOMM is also insensitive to neurocognitive deficits and has been validated in the detection of attempts to malinger neurocognitive symptoms (Tombaugh, 2002). The TOMM yields a total score; the proportion of items answered correctly is computed and scores below the cut score indicate malingering. Coefficient α = .95; sensitivity = .48; and specificity = 1.00.

Letter Memory Test (LMT; Inman et al., 1998*)*. The LMT is a variable-alternative forced choice recognition test shown to be insensitive to true cognitive impairment yet sensitive to malingering of impairment. It is also computer-administered, and is comprised of 45 items consisting of series of consonants. The LMT was designed largely in response to the potential effects of attorney coaching in reducing the utility of existing tests. Whereas the TOMM and the VSVT implement one "decoy" stimulus per item, the LMT varies in apparent difficulty by increasing both stimulus length (number of consonants) and the number of incorrect alternatives presented. The LMT has been validated in analogue studies with general college undergraduates, head injured undergraduates, head injured patients, and probable feigners in a forensic setting (Inman & Berry, 2002; Inman et al., 1998; Vickery et al., 2004). The proportion of items answered correctly is computed and scores below the cut score indicate malingering. Coefficient α = .94; sensitivity = .83 to .95; specificity = .94 to 1.00; and hit rate = .89.

Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997*)*. The SIMS is a 75-item true/false self-report measure consisting of five non-overlapping scales (Psychosis, Neurologic Impairment, Affective Disorders, Amnestic Disorders, and Low Intelligence) of 15 items each. Each scale is intended to assess feigning of a particular type of symptomatology. The Total score is simply the sum of the five scale scores and is used to determine the overall classification (malingering or not). The SIMS has been validated with college student simulators and male criminal defendants. Psychometric characteristics and proposed cut scores were reviewed previously.

Procedure

Participants were assessed over two days by the staff of the psychiatric practice. Staff members administering the various assessment instruments held Master's degrees in Psychology and were supervised by Ph. D.-level psychologists. All participants signed a consent form agreeing to allow their results to be used in this study. Per the practice's standard procedure, assessment began with participant completion of a lengthy information-gathering questionnaire designed by the staff. Participants were then administered multiple measures, which included tests of neurocognitive functioning and/or tests of psychopathology. Some instruments were selected by staff according to the type(s) of problems claimed by the plaintiff (e.g., if neurocognitive deficits were claimed, tests designed to assess neurocognitive functioning as well as feigning were administered). Other measures were administered to all participants. Individuals were included in the study if they provided signed consent to participate and completed the SIRS, the VSVT, the TOMM, the LMT, and the SIMS.

Results from this study were intended to assess the ability of the SIMS to screen for the two major types of feigning: psychiatric and neurocognitive. Toward that end, four pairs of groups were formed to allow for four sets of contrasts between malingering and honest respondents. Following are detailed descriptions of how each pair of contrast groups was formed along with rationales for each formulation.

Group Contrast #1: Psychiatric Symptom Malingering vs. Honest Reporting. The first contrast pair was designed to compare the performances of participants who malingered psychiatric symptoms, as determined by the SIRS, with those who were determined to have responded honestly on the SIRS. Participants were included in the Psychiatric – Malingering (PM) group if their SIRS results met at least one of the following criteria: (1) at least one SIRS Primary scale indicated Definite Malingering; (2) at least three SIRS Primary scales indicated Probable Malingering; or (3) the Total SIRS raw score exceeded 76. Participants were included in the Psychiatric – Honest (PH) group if they did not meet criteria for the PM group *and* at least six SIRS Primary scales indicated Honest responding. All remaining participants were classified as Psychiatric – Indeterminate (PI). (As their statuses could not be confidently determined, those with Indeterminate classifications were omitted from further analyses in all group contrasts.)

Table 2.1 displays the numbers of participants in the malingering, honest, and indeterminate groups along with the base rates of malingering for each of the four contrasts (the remaining three contrasts are defined below). As this table indicates, the PM group numbered 23, while the PH group included 172 subjects; the base rate of psychiatric malingering in this sample is 7.5%.

Comparisons between the PM and PH groups allow the evaluation of differences between litigants who malinger psychiatric symptoms and those who do not; specifically, the performance of the SIMS in distinguishing between these individuals can be assessed. As discussed earlier, a primary purpose of the SIMS is to screen out a substantial number of individuals who are clearly honest in their reports of psychiatric symptomatology.

Group Contrast #2: Neurocognitive Symptom Malingering vs. Honest Reporting. The second pair of groups allowed for the comparison of results of participants who malingered neurocognitive impairment with those of subjects who responded honestly on tests of neurocognitive effort. Participants were included in the Neurocognitive – Malingering (NM) group if the results of two or more tests among the VSVT, TOMM, and LMT indicated Malingering. Participants were classified as members of the Neurocognitive – Honest (NH) group if none of the three tests of neurocognitive effort (VSVT, TOMM, and LMT) indicated Malingering. All remaining participants were classified as Neurocognitive – Indeterminate (NI) and excluded from this comparison. Seventy-five subjects met criteria for the NM group; the NH group included 178, yielding a base rate of neurocognitive malingering equal to 24.4% (see Table 2.1).

Comparisons between these two groups make it possible to evaluate differences between litigants who malinger neurocognitive symptoms and those who do not; ultimately, the performance of the SIMS in distinguishing between these individuals can be assessed. As with the first contrast described above, this contrast represents a common real-world situation in which the SIMS is designed to be useful in separating individuals who are clearly honest in their neurocognitive symptom reports from those who are not.

Group Contrast #3: Psychiatric ***and*** *Neurocognitive Symptom Malingering vs. Honest Reporting.* The third contrast pair assessed a more blatant feigning strategy. Participants who met the criteria (as defined above) for *both* the PM group and the NM group were included in the Psychiatric and Neurocognitive – Malingering (PNM) group. Those who met criteria for *both* the PH and NH groups comprised the Psychiatric + Neurocognitive – Honest (PNH) group. All remaining participants were classified as Psychiatric/Neurocognitive – Indeterminate (PNI) and excluded from this contrast. The PNM group included 13 subjects, and 121 were members of the PNH group, resulting in a base rate of 4.2% for malingering of both psychiatric and neurocognitive symptoms (see Table 2.1).

Comparisons between these two groups allow the evaluation of differences between litigants who malinger both psychiatric *and* neurocognitive symptoms and those who do *not* malinger *either* type of symptom; again, the performance of the SIMS in distinguishing between these individuals can be tested. These analyses will be informative about litigants who malinger a wider range of symptoms, those who are most clearly responding honestly across a range of symptom reports, and how the contrasts between these individuals differ from more standard comparisons. It is expected that the distinctions between these two groups will be more dramatic, as members of each are held to a more stringent standard of malingering or honest responding. Individuals who malinger both types of symptoms may be thought of as malingering more "extremely".

Group Contrast #4: Psychiatric ***or*** *Neurocognitive Symptom Malingering vs. Honest Reporting.* The final set of groups allowed for a more general comparison between malingerers and honest responders. Any participant who qualified for the PM group *and/or* the NM group – that is, anyone who malingered psychiatric and/or neurocognitive symptoms – was included in the Any – Malingering (AM) group. The same PNH group of 121 subjects used in the third set of comparisons – those who responded honestly on the SIRS, the TOMM, the VSVT, *and* the LMT – formed the Honest group for this contrast as well. All remaining participants were classified as Any – Indeterminate (AI) and excluded from these comparisons. Eighty-five subjects belonged to the AM group, producing a base rate of 27.6% (refer to Table 2.1). Note that because the PM and NM groups overlapped substantially (sharing 13 participants), and the PH and NH groups overlapped even more so (with 121 in common), the base rate of psychiatric *or* neurocognitive malingering cannot be computed by adding the psychiatric base rate to the neurocognitive base rate. Figure 2.1 helps to clarify group breakdowns by depicting the overlap between the PM and NM groups, the PH and NH groups, and the PI and NI groups.

Comparisons between these final two groups allow the evaluation of differences between litigants who malinger psychiatric symptoms, neurocognitive symptoms, or both and those who do not malinger *any* of these symptoms; in particular, the performance of the SIMS in distinguishing between these individuals can be assessed. Due to the wider variation within the Any – Malingering group, these AM and PNH groups are expected to show less disparity than the first three pairs of groups, providing the most general yet perhaps most challenging test of the SIMS. In the "real world", the most likely scenario may be simply trying to identify who is malingering *in some way*. If the SIMS can reliably distinguish malingerers of any symptom from honest responders, it may demonstrate its most general utility.

Power Analyses

Previous research by Lewis et al. (2002) with a criminal forensic sample reported effect sizes for discrimination by the SIMS scales ranging from 1.1 for the Low Intelligence Scale score to 3.0 for the Total score. Table 2.2 shows power analyses based on the current sample of 308 participants and an $\alpha = .01$; as sample size varies across the four group contrasts, power levels for the various contrast groups range from 79.02% to 100% to detect effects of size 1.0.

Table 2.1

Group Ns and Malingering Base Rates for Four Contrasts

Type of Malingering	M†	H†	I†	Base Rate*
Psychiatric	23	172	113	**7.5%**
Neurocognitive	75	178	55	**24.4%**
Both Psychiatric and Neurocognitive	13	121	174	**4.2%**
Any	85	121	102	**27.6%**

†M = Malingering, H = Honest, I = Indeterminate

*N = 308

Table 2.2

*Power of Contrasts to Detect Effects of Various Sizes**

	Group Ns		Effect Size					
	M‡	H‡	0.2	0.5	0.8	1.0	2.0	3.0
PM-PH†	23	172	4.65%	36.58%	84.05%	97.06%	100.00%	100.00%
NM-NH†	75	178	12.87%	84.90%	99.93%	100.00%	100.00%	100.00%
PNM-PNH†	13	121	2.93%	18.83%	55.20%	79.02%	100.00%	100.00%
AM-PNH†	85	121	12.02%	82.34%	99.88%	100.00%	100.00%	100.00%

*α = .01

‡M = Malingering, H = Honest

†PM = Psychiatric – Malingering, PH = Psychiatric – Honest, NM = Neurocognitive – Malingering, NH = Neurocognitive – Honest, PNM = Psychiatric/Neurocognitive – Malingering, PNH = Psychiatric/Neurocognitive – Honest, AM = Any – Malingering

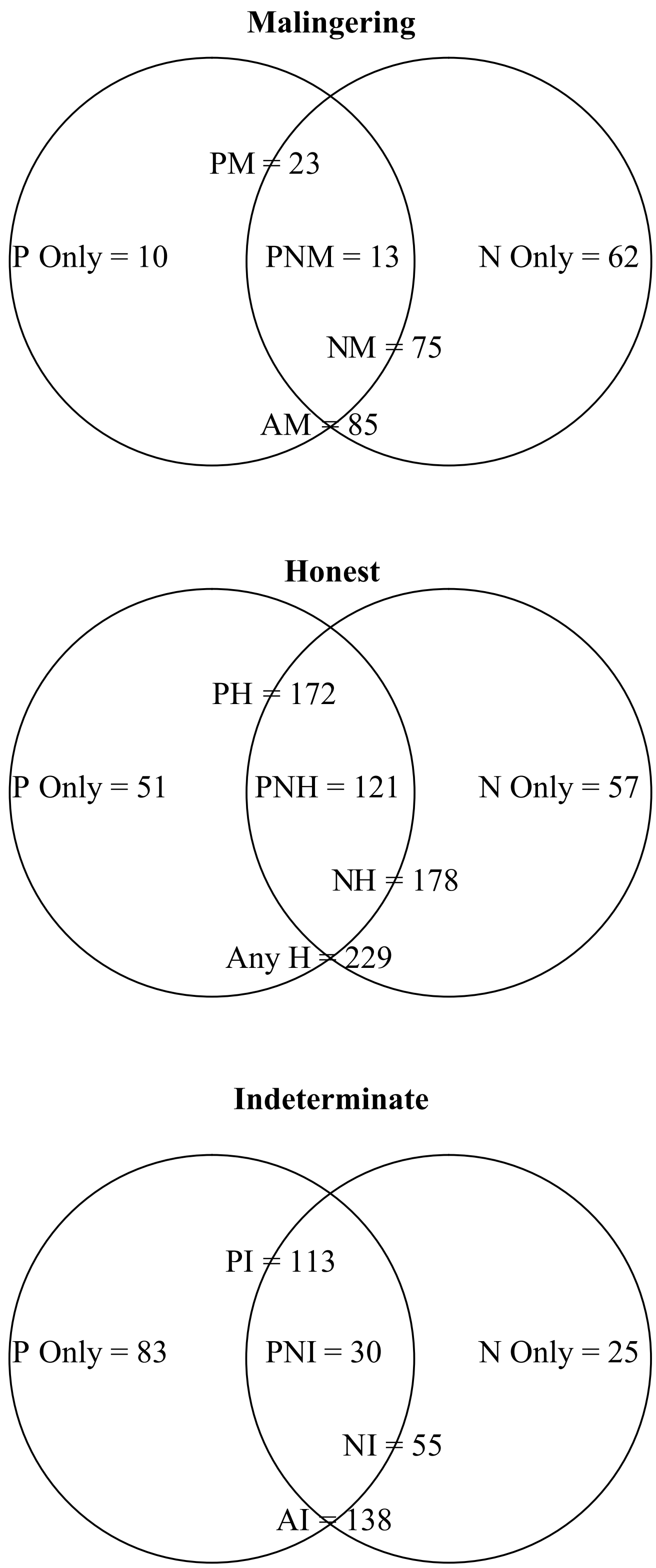

Figure 2.1. *Overlap between Psychiatric and Neurocognitive types of Malingering, Honest, and Indeterminate participants*

Chapter Three

Results

Analyses are presented below in three major sections. In the *Group Differences* section, results of t-tests and χ^2 tests on background variables and test scores are reported. Next, area under the receiver operating characteristic curve (AUC) computations and results of linear regression analyses illustrate the ability of the SIMS to predict malingering at the group level; these results are described in the *SIMS' Prediction of Malingering* section. Finally, in the *Base Rates and Classification Accuracy Statistics* section, the base rate and the SIMS' hit rate, sensitivity, specificity, positive predictive power, negative predictive power, incremental positive predictive power (IPPP, which equals PPP – Base Rate), and incremental negative predictive power (INPP, which equals NPP – [1 – Base Rate]) for each pair of contrast groups at each proposed SIMS cut score are presented.

Group Differences

Background variables, scores on criterion tests (the SIRS, TOMM, VSVT, and LMT), MMPI-2 scores, and performance on the SIMS by members of the Malingering and Honest groups were compared by way of t-tests and χ^2 tests for each pair of contrast groups. Due to the high number of statistical tests performed, an α level of .01 was applied to all tests of statistical significance. Additionally, effect sizes were computed using Cohen's *d* (Cohen, 1977).

Psychiatric Malingering vs. Honest Groups (PM vs. PH)

All participants who were determined to be malingering according to the SIRS – 23 in all – were included in the Psychiatric – Malingering (PM) group in this comparison. Those who were found to be honest according to the SIRS (n = 172) were included in the Psychiatric – Honest (PH) group. (See the Methods – *Measures* sub-section to review the SIRS Malingering and Honest criteria.) The base rate of malingering in this subsample was 7.5% (refer to Table 2.1).

Background variables. Table 3.1 displays means, standard deviations, t values, and effect sizes for age and education level. No significant differences between psychiatric malingerers and honest responders were detected. Proportions of participants in each category of race, gender, marital status, and type of claim (pain, psychiatric, brain damage, physical, and medical) are shown in Table 3.2. No significant differences between psychiatric malingerers and honest responders with respect to any of these variables were detected.

Neurocognitive Test Scores. Table 3.3 shows results for the tests of neurocognitive effort. Lower scores on these tests are suggestive of malingering. Interestingly, these psychiatric malingerers scored significantly *lower* than honest respondents on Trial 2 of the TOMM, the VSVT Difficult and Total, and the LMT, although neither PM nor PH group members were selected on the basis of neurocognitive malingering test scores. Effect sizes for differences on the neurocognitive malingering tests ranged from -1.01 (VSVT Difficult) to -1.25 (LMT). Further analysis provided a partial explanation for this effect, showing that 56.5% (13) of the PM group (n = 23) also malingered neurocognitive symptoms, while 70.3% (121) of the PH group (n = 172) were also honest responders with respect to the neurocognitive testing. (Relationships among psychiatric malingering classifications and neurocognitive feigning designations are clarified by examination of Table 3.4, which shows the numbers of individuals classified as malingering, honest, and indeterminate with respect to neurocognitive as well as psychiatric symptom reporting. Psychiatric designations are crossed with neurocognitive ones to show how these classifications intersect.)

MMPI-2 Scores. Means, standard deviations, t-test results, and *d* scores for all MMPI-2 Validity scales are presented in Table 3.5. Scores on all MMPI-2 Infrequency scales (F, Fb, F[p], and the F – K index) were higher in the malingering group; effect sizes ranged from 1.23 (F[p]) to 2.92 (Fb). These results are consistent with expectations, as the Infrequency scales were designed to be sensitive to endorsement of rare symptoms, and have seen some use in the detection of malingering (see MMPI-2 discussion in the Methods – *Measures* sub-section). Malingerers scored *lower* than honest respondents on the MMPI-2 K and S scales, with respective effect sizes of -0.92 and -1.10, suggesting lower levels of defensiveness in psychiatric malingerers, as might be predicted. The VRIN, TRIN, and L scales showed no significant differences between the groups.

As seen in Table 3.6, MMPI-2 clinical scale scores (except for scale 5) were higher in the malingering group than in the honest group. Effect sizes ranged from 0.69 (scale 9) to 2.03 (scale 6). This suggests that, as one might anticipate, malingerers endorsed symptoms across a range of psychopathology at a significantly higher-than-normal rate.

SIMS Scores. As shown in Table 3.7, all SIMS scores were significantly higher in the malingering group. Robust effect sizes were produced: 2.55 (Total), 1.85 (N), 1.42 (Af), 3.27 (P), 0.94 (LI), and 1.77 (Am). This provides initial confirmation that the SIMS Total and all its sub-scales show some ability to distinguish groups malingering psychiatric symptoms from those who are not. The Psychotic symptoms (P) scale showed the largest effect, suggesting that the effect of psychiatric malingering on SIMS scores in this sample may be largely due to endorsement of psychotic symptoms. This is consistent with the largest SIRS effect size for the BL (Blatant symptoms) scale ($d = 7.21$) and the largest MMPI-2 clinical scale effect sizes for scales 6 (Paranoia) and 8 (Schizophrenia) of 2.03 and 1.97, respectively (see Table 3.6).

Neurocognitive Malingering vs. Honest Groups (NM vs. NH)

All participants who were determined to be malingering according to at least two of the three tests of neurocognitive effort (the TOMM, VSVT, and LMT; n = 75) were included in the Neurocognitive – Malingering (NM) group in this comparison. Those who were found to be honest according to the TOMM, VSVT, and LMT were included in the Honest (NH) group (n = 178). (See the Methods – *Measures* sub-section to review the TOMM, VSVT, and LMT Malingering and Honest criteria.) The base rate of malingering in this subsample was 24.4% (see Table 2.1).

Background variables. Table 3.8 displays results of t-tests and effect sizes for age and education level for these two groups. As shown, no significant differences were detected. In Table 3.9, χ^2 results for comparisons on race, gender, marital status, and claim types are presented, and, again, no differences were found between neurocognitive malingerers and honest responders.

SIRS Scores. Notably, as seen in Table 3.10, all SIRS scales and the SIRS Total showed substantially *higher* scores in the neurocognitive malingering group, suggestive of psychiatric malingering by the NM participants; effect sizes ranged from .49 on the RS (Rare Symptoms) scale to .99 on the Total score. This is interesting because psychiatric symptom malingering status was not a criterion for NM or NH group membership. While 68% (121) of the NH group members (n = 178) were also honest with respect to psychiatric malingering, only 17.3% (13) of the NM group members (n = 75) also malingered psychiatric symptoms (see Table 3.4). These results suggest that the SIRS may offer some utility in discriminating neurocognitive malingerers from honest responders. It is interesting that the neurocognitive malingerers (in comparison to neurocognitive honest responders) showed a greater effect size – .91 – for the SU (Subtle Symptoms) scale than for any other sub-scale, reflecting a pattern of differential psychiatric symptom endorsement distinct from that of the psychiatric malingerers (in comparison with the psychiatric honest responders), who, instead, showed the greatest effect size for Blatant Symptom reporting.

MMPI-2 Scores. Table 3.11 displays means, standard deviations, t-test results, and *d* scores for all MMPI-2 Validity scales. Scores on all MMPI-2 Infrequency scales (F, Fb, F[p], and the F – K index) were higher in the neurocognitive malingering group; effect sizes ranged from 0.57 (F[p]) to 1.16 (F). This is particularly notable given that these scales reflect endorsement of infrequently reported *psychiatric* symptoms. No differences in VRIN, TRIN, L, K, or S scores were found.

As seen in Table 3.12, MMPI-2 clinical scale scores (except for scales 4, 5, and 9) were higher in the malingering group than in the honest group. Effect sizes ranged from 0.47 (scale 1) to 1.12 (scale 8). These results, as those from the SIRS (above), suggest that neurocognitive malingerers are also reporting significantly more symptoms of various types of psychopathology. It is interesting that the largest effect size among the MMPI-2 clinical scales is for scale 8 (Schizophrenia), suggesting (unlike the SIRS results) that the largest differences were in the reporting of psychotic-like symptoms.

SIMS Scores. Table 3.13 shows that all SIMS scores were higher for neurocognitive malingerers, with effect sizes of 1.13 (Total), 0.80 (N), 0.61 (Af), 0.82 (P), 0.41 (LI), and 1.23 (Am). Given the limited overlap of the PM group into the NM group (only 13 of 75, or 17.3%, of NM group members are also PM group members – see Table 3.4), it is noteworthy that the SIMS Total and all its sub-scales show significant effectiveness in distinguishing individuals malingering neurocognitive symptoms from those who are not. The largest effect size for the Am (Amnestic symptoms) scale ($d = 1.23$) is not surprising for a group defined by its poor performance on neurocognitive malingering measures that appear to be memory tests. The N

(Neurologic symptom) scale *d* of .80 is in the mid-range of the SIMS scale effect sizes. At the same time, it is interesting that the LI (Low Intelligence) scale effect size of .41 is the smallest, suggesting that there is more of a difference between neurocognitive malingerers and honest responders in the endorsement of Affective and Psychotic symptoms than there is in Low Intelligence symptom reporting.

Both Psychiatric and Neurocognitive Malingering vs. Honest Groups (PNM vs. PNH)

Participants who were determined to be malingering according to the SIRS *and* at least two of the three tests of neurocognitive effort (the TOMM, VSVT, and LMT; n = 13) were included in the Malingering (PNM) group for this comparison. Those who were found to be honest according to the TOMM, VSVT, LMT, *and* the SIRS (n = 121) were included in the Honest (PNH) group. The base rate of malingering in this subsample was 4.2% (refer to Table 2.1).

Background variables. Table 3.14 shows results of t-tests for age and education level. No significant differences between malingerers of both psychiatric and neurocognitive symptoms and honest responders were found. Additionally, Table 3.15 shows that no differences in the rates of any race, gender, marital status, or claim type were detected.

MMPI-2 Scores. Results of t-tests and effect sizes for all MMPI-2 Validity scales are shown in Table 3.16. Scores on all MMPI-2 Infrequency scales (F, Fb, F[p], and the F – K index) were higher in the malingering group, suggesting more psychiatric symptom endorsement; effect sizes ranged from 1.44 (F[p]) to 3.21 (Fb). In addition, psychiatric + neurocognitive malingerers scored *lower* than honest respondents on the MMPI-2 K (Defensiveness) and S (Superiority) scales, with respective effect sizes of -0.77 and -0.91. As discussed in previous sections, these results make sense given the respective purposes of these MMPI-2 scales.

Further, MMPI-2 clinical scale scores (except for scales 5 and 9), as reported in Table 3.17, were *higher* in the malingering group than in the honest group. Effect sizes ranged from 1.05 (scales 3 and 0) to 2.67 (scale 8). Again, malingerers – all of whom were determined to have malingered psychiatric symptoms – endorsed a variety of symptoms of psychopathology.

SIMS Scores. Participants who malingered both psychiatric and neurocognitive symptoms scored significantly higher on all SIMS scales as well. Table 3.18 shows these results, which yielded effect sizes of 3.37 (Total), 2.20 (N), 1.45 (Af), 4.25 (P), 1.28 (LI), and 2.67 (Am). As expected, the effect sizes are higher for this contrast than for the first two. These PNM malingerers are feigning both psychiatric problems *and* neurocognitive deficits – and the PNH group members are honest with respect to reporting of both types of symptoms. Thus, it seems logical that the differences between these two groups would be starker. Interestingly, the largest differences between these malingering and honest groups were demonstrated to be in the endorsement of severe symptoms (SIRS SEV scale, d = 12.47), schizophrenia-related symptoms (MMPI-2 Scale 8, d = 2.67), and psychotic symptoms (SIMS P scale, d = 4.25).

Any Malingering vs. Honest Groups (AM vs. PNH)

The final set of contrasts compared malingerers of *any* type of symptoms (psychiatric and/or neurocognitive) with the PNH group. Participants were included in the Malingering (AM) group if they were determined to be malingering according to *one or both* of the following criteria: (a) the SIRS; and/or (b) at least two of the three tests of neurocognitive effort (the TOMM, VSVT, and LMT). Eighty-five participants met this condition. As in the previous contrast, those who were found to be honest according to the TOMM, VSVT, LMT, *and* the SIRS – n = 121 – were included in the Honest (PNH) group. The base rate of malingering in this subsample was 27.6% (see Table 2.1).

Background variables. Table 3.19 shows that no significant differences in age or education level were found when comparing the AM and PNH groups. In addition, the results in Table 3.20 indicate that no significant differences with respect to race, gender, marital status, or claim type were found between malingerers of psychiatric *or* neurocognitive symptoms and honest responders.

MMPI-2 Scores. Results of t-tests and effect sizes for all MMPI-2 Validity scales are presented in Table 3.21. Scores on all MMPI-2 Infrequency scales (F, Fb, F[p], and the F – K index) were, again, higher in the malingering group; effect sizes ranged from 0.71 (F[p]) to 1.59 (F). Malingerers scored *lower* than honest respondents on the MMPI-2 K and S scales (with respective effect sizes of -0.50 and -0.45). These results are expected for reasons already discussed; and due to the more inclusive AM group, the effect sizes are smaller than in previous comparisons.

MMPI-2 Clinical scale t-test results and *d* values appear in Table 3.22. Clinical scale scores (except for scales 4, 5, and 9) were higher in the malingering group than in the honest group. Effect sizes ranged from 0.73 (scale 3) to 1.46 (scale 8). Even in this "weaker" contrast, a wide range of psychopathology is reported and effect sizes remain large.

SIMS Scores. Finally, comparison of participants malingering either psychiatric or neurocognitive symptoms or both with those malingering neither type of symptom shows higher scores for malingerers on all SIMS scales (see Table 3.23). Effect sizes produced are: 1.46 (Total), 1.14 (N), 0.97 (Af), 0.94 (P), 0.47 (LI), 1.39 (Am). All but one (LI) of these is large, and all are smaller than effects for comparisons involving only psychiatric symptom malingerers, but larger than those for the NM-NH comparison. Once again, preliminary support for the screening utility of the SIMS at the group level is suggested.

As one might predict, there is more variability in the types of symptoms tapped by the scales generating the largest effect sizes in this AM vs. PNH comparison. The SIRS scale showing the greatest effect size is SEV (Severity of symptoms; $d = 2.04$); the largest effect size for an MMPI-2 clinical scale is that for scale 8 (Schizophrenia; $d = 1.46$); but the most substantial SIMS scale effect size is the Total ($d = 146$), followed closely by Am (Amnestic symptoms; $d = 1.39$) and then N (Neurologic symptoms; $d = 1.14$). Although the PM vs. PH and NM vs. NH contrasts suggested substantial overlap between PM and NM groups in the type of symptom endorsement, only 27.1% (23) of the AM group (n = 85) was determined to have malingered psychiatric symptoms; it may then stand to reason that the SIMS would show smaller effect sizes for psychiatric symptom scales than for the Am and N scales in this comparison.

Summary of Group Differences on Background Variables

No significant differences with respect to race, gender, age, educational level, or marital status were found between any of the contrast group pairs. Additionally, no differences in the rates of any types of claims (pain, psychiatric, brain damage, physical, or medical) were detected in any of these comparisons. These results indicate that malingering and honest groups were comparable on these demographic and litigation variables.

Summary of Group Differences on Test Scores

All differences found on SIRS (Total and sub-scale), TOMM, VSVT, LMT, and MMPI-2 Validity and Clinical scale scores were in the expected directions. Further, in *each* comparison, the SIMS sub-scale and Total scores were significantly higher in the malingering group, lending general support to the role of the SIMS scales as indicators of malingering, at least at the group level.

Examination of effect sizes for the various psychiatric symptom-oriented test score differences across all four contrasts shows a clear pattern of results. Table 3.24 compares effect sizes of some of the test score differences for the four pairs of malingering and honest groups and highlights this pattern. For most of the scores, including those for all SIRS scales, the SIRS Total, MMPI-2 Infrequency scales, MMPI-2 clinical scales, all SIMS scales, and the SIMS Total, the PNM – PNH comparisons yielded the largest effect sizes. This is consistent with predictions that the greatest disparities would be found between the PNM and PNH groups. These malingerers *do* appear to malinger more "extremely" in comparison to participants who are clearly honest in all symptom reporting, as contrasted with malingerers in the other comparisons. The best support for this hypothesis in this group comparison is provided by the MMPI-2 and SIMS results, as scores on these tests were not included in the criteria for group membership.

The pattern continues with all these same tests showing the next-largest effect sizes in the PM – PH comparisons. These results make sense for the SIRS, which was designed to detect psychiatric symptom malingering, the MMPI-2 clinical scales, which measure psychiatric symptom endorsement, and the MMPI-2 Infrequency scales, which are intended to detect high endorsement of infrequently reported psychiatric symptoms. It is interesting that all of the SIMS scales and the Total are also next most successful at detecting group differences between psychiatric malingerers and honest responders (see Table 3.25 for all SIMS subscale and Total effect sizes for all comparisons).

Among these same tests, the AM – PNH comparisons produced the third highest set of effect sizes. As discussed previously, the weaker contrast between these two groups was expected to present more of a challenge to tests in attempts to distinguish malingerers from honest responders. Notably, however, at this group level of analysis, the SIMS Total *d* score remained substantial (1.46).

Finally, the smallest effect sizes in these test score differences were found for the NM – NH comparisons. For the SIRS and MMPI-2 scales, these results make sense, as these tests are designed to detect reporting of false *psychiatric* symptoms. It is interesting, however, that not only the SIMS Total, but *all* SIMS scales, including those designed to detect malingering reports of Low Intelligence, Amnestic, and Neurologic symptoms, show the smallest effect sizes in this NM – NH comparison (refer again to Table 3.25).

For an overall picture of SIMS results, Table 3.25 shows the effect sizes for all SIMS scales and the SIMS Total across all four sets of contrasts. LI effect sizes are consistently smallest, whereas values for the P scale are highest for both the Psychiatric and Both comparisons; the Am effect size is largest for the Neurocognitive comparison; and the SIMS Total produced the greatest effect in the Any comparison.

The pattern found in the effect sizes for differences in *neurocognitive* test scores, however, is different. Returning to Table 3.24, it can be seen that for the TOMM Trial 2, VSVT Difficult, VSVT Total, and LMT, the positions of the NM – NH results and the PM – PH results were reversed – the order of comparisons from largest to smallest effect sizes is: Both Psychiatric & Neurocognitive > Neurocognitive > Any > Psychiatric. This result is certainly consistent with the purposes of these tests and, therefore, with expectations. Tests designed to detect neurocognitive symptom feigning (and not designed to identify psychiatric malingerers) should perform better at distinguishing neurocognitive malingerers from honest responders than they do at distinguishing psychiatric malingerers from honest responders.

SIMS' Prediction of Malingering

Areas Under the Curve

Areas Under the Curve (AUCs) were computed for the performance of the SIMS Total score in distinguishing malingerers from honest responders in each of the four comparisons. AUC values, all significant at $p < .001$, are: .899 (Psychiatric; SE = .026), .755 (Neurocognitive; SE = .035), .983 (Both; SE = .010), and .822 (Any; SE = .031). From largest to smallest, these AUCs follow the same pattern as SIMS scores: Both > Psychiatric > Any > Neurocognitive. Such considerable AUC magnitudes indicate that the SIMS Total score has substantial potential utility in screening for malingerers defined according to any of the four methods employed in this study.

Regression Analyses

Linear regression analyses were conducted in order to examine the relative contributions of the SIMS Total score and the MMPI-2 Infrequency scale scores in the prediction of criterion scale scores. First, the incremental utility of the Infrequency scales over that of the SIMS Total score was computed for each criterion measure; then the utility of the SIMS Total score in predicting each criterion score over and above that of the Infrequency scales was evaluated.

Specifically, conditional stepwise regressions of each criterion measure (the SIRS Total, TOMM T2, VSVT Difficult, VSVT Total, and LMT) were performed onto the SIMS Total, entered at Step 1, and the MMPI-2 Infrequency scales (F, Fb, and F[p]), which were entered in subsequent Steps. As can be seen in Table 3.26, the scales F and Fb (but not F[p]) provided incremental utility in the prediction of the SIRS Total, and F alone added significantly to the prediction of the VSVT Difficult, VSVT Total, and LMT scores. Regression of the TOMM T2 showed no increment in prediction over that of the SIMS Total by any Infrequency scales.

More importantly, in subsequent regressions, the Infrequency scales were conditionally entered in Step 1, and the SIMS Total was entered in Step 2. The incremental utility of the SIMS Total in predicting each malingering scale score over and above the level of prediction yielded by the Infrequency scales was demonstrated in each of the five regressions ($p \leq .001$ for all; refer to Table 3.26).

Base Rates and Classification Accuracy Statistics

Ultimately, the utility of a malingering instrument in a forensic application is assessed by examining its operating characteristics when used to classify individuals as either malingering or honest. In the case of a screen, such as the SIMS, its ability to capture malingerers is primary; in an attempt to ensure that all malingerers are properly identified, a moderate rate of false positives is tolerable, although a screening instrument is useful only if it does, in fact, screen out some significant proportion of honest respondents.

Further, although sensitivity and specificity statistics will allow computation of risks of false positives and false negatives, only positive predictive power and negative predictive power are informative about the likelihood that a particular individual whose test sign has been determined is in fact malingering or honest. Therefore, these are the statistics that are truly useful to the forensic examiner. Further, the clinical utility of any such test is not fully appreciated without comparing its PPP and NPP to the classification accuracy of random assignment – i.e., *incremental* positive and negative predictive power are most informative. The base rate of malingering must be known in order to compute these additional figures.

The base rate of each operational definition of malingering was computed based on the participants' statuses on the criterion measures as described in the *Procedures* section. For each pair of groups, sensitivity, specificity, overall classification accuracy (hit rate), positive predictive power, negative predictive power, incremental positive predictive power, and incremental negative predictive power were computed for the SIMS Total score using cut scores of greater than 14 and greater than 16, as suggested in previous literature, in an attempt to cross-validate these cut scores.

Cut Scores Recommended in Previous Studies

Table 3.27 shows the classification accuracy statistics for all four sets of comparisons using a cut score of > 14, while Table 3.28 displays the same information for a cut score of > 16. Each of these tables also includes similar statistics from previous studies that reported results at the respective cut score.

As can be seen in Table 3.27, sensitivity values using a cut score of > 14 to detect psychiatric or "psychiatric and neurocognitive" malingering are high (.957 and 1.000, respectively) and comparable to rates found in studies by Edens et al. (1999) and Smith & Burger (1997; .960 in each). Detection of neurocognitive or "any" malingering produced lower sensitivity rates of .800 and .812, respectively. NPP values are .989 and 1.000 for the prediction of honest responding with respect to psychiatric and "both" types of symptoms, respectively, compared with .958 and .786 in Edens et al. (1999) and Smith & Burger (1997), respectively. Again, NPP values for honest responding to neurocognitive or "any" symptoms were lower: .848 and .812, respectively. These results support the utility of this cut score in this sample for detecting psychiatric malingerers, but suggest that it is notably less effective at detecting neurocognitive malingering than psychiatric malingering, as most malingerers in the "any" malingering group (75 of 85) were determined to be feigning neurocognitive symptoms only.

While specificity and PPP are less important for any instrument designed to function as a screening measure, it is worth noting that other researchers have found much higher specificity at a cut score of > 14 than shown in this study. This, along with much higher base rates, contributes to the much higher PPP reported in previous studies (also shown in Table 3.27).

Table 3.28 displays statistics produced using a SIMS Total cut score of > 16. Sensitivity values for the detection of psychiatric or "psychiatric and neurocognitive" malingering are high (.957 and 1.000, respectively). These are comparable to Lewis et al.'s (2002) finding of 1.000. (Smith, 1992 and Rogers et al., 1996 did not report specificity figures.) Detection of neurocognitive and "any" malingering produced lower sensitivity rates of .747 and .765, respectively. NPP values are .991 and 1.000 for the prediction of honest responding with respect to psychiatric and "both" types of symptoms, respectively, compared with 1.000 and .620 in Lewis et al. (2002) and Rogers et al. (1996), respectively. Again, NPP values for honest responding to neurocognitive and "any" type of symptoms were somewhat lower: .849 and .810, respectively. These results are supportive of the utility of a cut score of > 16 in this sample for screening for psychiatric malingerers; however, as with the cut score of > 14, this cut score is appreciably less effective against neurocognitive malingering than psychiatric malingering in this sample.

Table 3.28 shows that samples in previous studies had higher base rates than were identified in the present study, along with correspondingly higher PPP values; however, the authors of only one of the three previous studies (Lewis et al., 2002) reported a specificity figure for the cut score of > 16, and this value (54.0%) was comparable to those found in the current study. Due to differences in specificity, hit rates at both > 14 and > 16 were also generally higher in the previous SIMS studies than those determined herein (see Tables 30 and 31).

To summarize, then, the cut scores of > 14 and > 16 both show quite good sensitivity and negative predictive power in this sample when the SIMS is used to screen for malingerers of psychiatric or "both psychiatric and neurocognitive" symptoms, all of whom are malingering psychiatric symptomatology. However, these cut scores perform less than optimally when screening for neurocognitive or "psychiatric and/or neurocognitive" malingering (sensitivity values range from only .747 to .812, and at the base rates in this sample, these figures yield moderate NPPs, ranging from .812 to .849). Consequently, the cut score that produces maximum sensitivity and NPP when screening for neurocognitive malingering has been determined for this sample. Classification accuracy statistics computed for this cut score are shown in Table 3.29.

Cut Score Maximizing Sensitivity and NPP for NM-NH and AM-PNH Comparisons

The cut score which generates the highest possible sensitivity and NPP (both equal to 100%) for both the NM-NH and AM-PNH comparisons is a SIMS Total score > 4. Corresponding specificities are rather low: 51.0% for NM-NH and 58.0% for AM-PNH. Resultant hit rates are also rather poor: 33.2% and 44.7% for NM-NH and AM-PNH, respectively. At such a low cut score, very few honest responders (9 of 178 in the NH group and 7 of 121 in the PNH group) are screened out.

Table 3.1

Group Comparisons of Continuous Background Variables

Psychiatric Malingering vs. Honest (PM vs. PH)

	Psychiatric Malingering Group†			Psychiatric Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
Age	23	39.04	9.777	172	42.92	12.096	-1.475	193	-0.33
Education level	23	11.78	1.536	171	12.2.7	2.724	-0.837	192	-0.19

† $N_{Mal} = 23$, $N_{Hon} = 172$, $N_{M+H} = 195$

*p < .01

Table 3.2

Group Comparisons of Discrete Background Variables

Psychiatric Malingering vs. Honest (PM vs. PH)

		Psychiatric Malingering†		Psychiatric Honest†			
		N	%	N	%	χ^2	df
Demographics							
Race	White	23	100.0%	157	91.3%	21.730	2
	Af-Am	0	0.0%	14	8.1%		
	Other	0	0.0%	1	0.6%		
Gender	Male	15	65.2%	122	70.9%	0.367	1
	Female	8	34.8%	49	28.5%		
Marital status	Single	5	21.7%	20	11.6%	4.535	4
	Married	12	52.2%	116	67.4%		
	Separated	0	0.0%	2	1.2%		
	Divorced	6	26.1%	26	15.1%		
	Widowed	0	0.0%	4	2.3%		
Claim Types							
Pain	Yes	5	21.7%	70	40.7%	2.921	1
Psychiatric	Yes	22	95.7%	124	72.1%	5.672	1
Brain damage	Yes	5	21.7%	72	41.9%	3.437	1
Physical	Yes	7	30.4%	40	23.3%	0.501	1
Medical	Yes	1	4.3%	6	3.5%	0.041	1

†$N_{Mal} = 23$, $N_{Hon} = 172$, $N_{M+H} = 195$

*$p < .01$

Table 3.3

Group Comparisons of Neurocognitive Test Scores

Psychiatric Malingering vs. Honest (PM vs. PH)

	Psychiatric Malingering Group†			Psychiatric Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
TOMM									
Trial 2	23	84.26	20.166	172	96.53	9.718	-2.875*	23.385	-1.12
VSVT									
Diff	23	61.91	29.985	172	84.34	21.170	-3.474*	25.016	-1.01
Total	23	76.52	19.730	172	91.29	12.576	-3.496*	24.446	-1.10
LMT	23	79.53	22.980	172	94.99	11.044	-3.178*	23.377	-1.25

†$N_{Mal} = 23$, $N_{Hon} = 172$, $N_{M+H} = 195$

*$p < .01$

Table 3.4

Psychiatric and Neurocognitive Malingering, Honest, and Indeterminate Classification Ns

	PH†	PI†	PM†	Total
NH†	121	50	7	**178**
NI†	22	30	3	**55**
NM†	29	33	13	**75**
Total	**172**	**113**	**23**	**308**

†PH = Psychiatric – Honest, PI = Psychiatric – Indeterminate, PM = Psychiatric – Malingering, NH = Neurocognitive – Honest, NI = Neurocognitive – Indeterminate, NM = Neurocognitive – Malingering

Table 3.5

Group Comparisons of MMPI-2 Validity Scale Scores

Psychiatric Malingering vs. Honest (PM vs. PH)

	Psychiatric Malingering Group†			Psychiatric Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
VRIN	23	48.30	8.314	169	51.46	8.761	-1.628	190.000	-0.36
TRIN	23	58.39	7.578	169	58.78	8.126	-0.217	190.000	-0.05
L	23	55.17	8.600	169	58.76	10.606	-1.554	190.000	-0.35
K	23	38.17	7.414	169	48.01	11.082	-5.569*	36.964	-0.92
S	23	36.52	10.655	169	48.36	10.827	-4.929*	190.000	-1.10
F	23	96.17	15.305	169	58.21	13.868	12.164*	190.000	2.70
Fb	23	111.65	11.578	169	61.90	17.722	17.945*	37.763	2.92
F(p)	23	64.74	13.676	169	50.60	11.209	5.523*	190.000	1.23
F – K	23	8.61	7.334	168	-6.57	10.171	6.909*	189.000	1.54

†$N_{Mal} = 23$, $N_{Hon} = 172$, $N_{M+H} = 195$

**p* < .01

Table 3.6

Group Comparisons of MMPI-2 Clinical Scale Scores

Psychiatric Malingering vs. Honest (PM vs. PH)

	Psychiatric Malingering Group†			Psychiatric Honest Group†					
	N	M	SD	N	M	SD	t	Df	*d*
MMPI-2									
1	22	92.86	7.376	168	77.66	12.580	8.227*	39.326	1.27
2	22	96.86	8.317	168	78.81	14.535	8.605*	40.345	1.30
3	22	90.55	9.231	168	77.48	15.700	5.655*	39.216	0.87
4	22	73.23	7.746	168	58.54	10.112	6.559*	188.000	1.49
5	22	51.32	7.473	168	46.76	8.365	2.430	188.000	0.55
6	22	88.91	17.979	168	57.87	14.965	8.930*	188.000	2.03
7	22	96.32	5.677	168	72.22	14.511	14.616*	66.214	1.78
8	22	97.95	20.151	168	66.71	15.362	8.630*	188.000	1.97
9	22	54.41	6.478	168	47.31	10.776	3.015*	188.000	0.69
0	22	77.23	9.995	168	62.80	13.088	6.119*	31.291	1.13

†$N_{Mal} = 23$, $N_{Hon} = 172$, $N_{M+H} = 195$

*$p < .01$

Table 3.7

Group Comparisons of SIMS Total and Subscale Scores

Psychiatric Malingering vs. Honest (PM vs. PH)

	Psychiatric Malingering Group†			Psychiatric Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
SIMS									
Total	23	31.57	8.851	172	14.41	6.466	11.398*	193.000	2.55
N	23	8.13	2.719	172	3.42	2.525	8.331*	193.000	1.85
Af	23	9.17	1.466	172	5.95	2.374	9.064*	39.498	1.42
P	23	3.35	3.157	172	0.31	0.643	4.605*	22.245	3.27
LI	23	3.70	1.964	172	2.15	1.601	4.226*	193.000	0.94
Am	23	7.17	3.172	172	2.53	2.544	7.975*	193.000	1.77

†$N_{Mal} = 23$, $N_{Hon} = 172$, $N_{M+H} = 195$

*$p < .01$

Table 3.8

Group Comparisons of Continuous Background Variables

Neurocognitive Malingering vs. Honest (NM vs. NH)

	Neurocognitive Malingering Group†			Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
Age	75	40.57	10.940	178	41.88	11.768	-0.821	251	-0.11
Education level	75	11.68	2.786	177	12.33	2.548	-1.794	250	-0.25

†$N_{Mal} = 75$, $N_{Hon} = 178$, $N_{M+H} = 253$

*$p < .01$

Table 3.9

Group Comparisons of Discrete Background Variables

Neurocognitive Malingering vs. Honest (NM vs. NH)

		Neurocognitive Malingering†		Neurocognitive Honest†			
		N	%	N	%	χ^2	df
Demographics							
Race	White	69	92.0%	164	92.1%	0.456	2
	Af-Am	6	8.0%	13	7.3%		
	Other	0	0.0%	1	0.6%		
Gender	Male	52	69.3%	118	66.3%	0.221	1
	Female	23	30.7%	60	33.7%		
Marital status	Single	9	12.0%	22	12.4%	2.743	4
	Married	48	64.0%	19	66.9%		
	Separated	1	1.3%	4	2.2%		
	Divorced	14	18.7%	28	15.7%		
	Widowed	3	4.0%	2	1.1%		
Claim Types							
Pain	Yes	34	45.3%	72	40.4%	0.442	1
Psychiatric	Yes	64	85.3%	138	77.5%	1.798	1
Brain damage	Yes	24	32.0%	62	34.8%	0.215	1
Physical	Yes	27	36.0%	43	24.2%	3.499	1
Medical	Yes	7	9.3%	6	3.4%	3.803	1

†$N_{Mal} = 75$, $N_{Hon} = 178$, $N_{M+H} = 253$

*$p < .01$

Table 3.10

Group Comparisons of SIRS Scores

Neurocognitive Malingering vs. Honest (NM vs. NH)

	Neurocognitive Malingering Group†			Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
SIRS									
Total	75	43.03	26.113	178	22.91	18.035	6.088*	104.962	0.99
RS	75	1.33	2.226	178	0.59	1.233	2.722*	93.713	0.49
SC	75	1.47	1.671	178	0.60	1.086	4.133*	101.326	0.69
IA	75	0.96	1.144	178	0.38	0.850	3.979*	109.901	0.62
BL	75	4.35	4.397	178	2.02	2.751	4.241*	99.296	0.72
SU	75	13.545	7.600	178	7.70	5.812	5.870*	112.115	0.91
SEL	75	11.83	5.463	178	7.38	4.777	6.475*	251.000	0.89
SEV	75	6.09	6.047	178	2.38	3.910	4.908*	101.051	0.82
RO	75	3.55	2.637	178	1.86	1.892	5.023*	107.463	0.80

†$N_{Mal} = 75$, $N_{Hon} = 178$, $N_{M+H} = 253$

**p* < .01

Table 3.11

Group Comparisons of MMPI-2 Validity Scale Scores

Neurocognitive Malingering vs. Honest (NM vs. NH)

	Neurocognitive Malingering Group†			Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
VRIN	74	50.82	9.772	173	52.07	8.598	-1.000	245.000	-0.14
TRIN	75	58.29	6.976	173	57.76	8.666	0.468	246.000	0.06
L	75	56.84	9.744	173	56.79	10.922	0.033	246.000	0.00
K	75	43.04	10.427	173	45.86	10.596	-1.935	246.000	-0.27
S	75	43.95	10.438	173	45.48	10.753	-1.040	246.000	-0.14
F	75	78.99	18.753	173	59.47	16.095	7.845*	123.400	1.16
Fb	75	88.08	23.595	173	66.95	22.155	6.764*	246.000	0.94
F(p)	75	57.47	12.341	173	50.71	11.789	4.086*	246.000	0.57
F – K	75	1.49	8.880	172	-5.25	10.580	4.827*	245.000	0.67

†$N_{Mal} = 75$, $N_{Hon} = 178$, $N_{M+H} = 253$

*$p < .01$

Table 3.12

Group Comparisons of MMPI-2 Clinical Scale Scores

Neurocognitive Malingering vs. Honest (NM vs. NH)

	Neurocognitive Malingering Group†			Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
1	75	86.83	9.964	172	77.50	13.425	6.056*	186.860	0.47
2	75	91.91	11.838	172	80.31	15.509	6.414*	182.084	0.81
3	75	86.53	12.989	172	77.05	16.353	4.864*	175.387	0.62
4	75	75.320	85.112	172	59.75	11.069	2.342	245.000	0.46
5	75	49.36	8.112	172	48.40	9.354	0.770	245.000	0.11
6	75	72.97	19.674	172	61.45	16.489	4.754*	245.000	0.66
7	75	86.92	14.665	172	73.92	14.499	6.458*	245.000	0.89
8	75	87.27	16.076	172	68.38	17.306	8.056*	245.000	1.12
9	75	49.75	11.064	172	48.99	10.737	0.502	245.000	0.07
0	75	73.19	12.029	171	63.58	12.544	5.595*	245.000	0.78

†$N_{Mal} = 75$, $N_{Hon} = 178$, $N_{M+H} = 253$

*$p < .01$

Table 3.13

Group Comparisons of SIMS Total and Subscale Scores

Neurocognitive Malingering vs. Honest (NM vs. NH)

	Neurocognitive Malingering†			Neurocognitive Honest†					
	N	M	SD	N	M	SD	t	df	*d*
SIMS									
Total	75	24.76	10.754	178	15.55	7.074	6.821*	102.028	1.13
N	75	6.21	3.633	178	3.75	2.828	5.235*	113.508	0.80
Af	75	7.88	2.205	178	6.44	2.419	4.424*	251.000	0.61
P	75	1.61	2.614	178	0.46	0.903	3.726*	81.535	0.82
LI	75	3.04	2.076	178	2.31	1.671	2.682*	116.313	0.41
Am	75	5.97	3.377	178	2.53	2.552	7.920*	111.189	1.23

†$N_{Mal} = 75$, $N_{Hon} = 178$, $N_{M+H} = 253$

*$p < .01$

Table 3.14

Group Comparisons of Continuous Background Variables

Both Psychiatric and Neurocognitive Malingering vs. Honest (PNM vs. PNH)

	Both Psychiatric & Neurocognitive Malingering Group†			Both Psychiatric & Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
Age	13	41.92	8.760	121	42.95	12.702	-0.382	17.942	-0.08
Education level	13	12.08	1.498	120	12.47	2.722	-0.507	131.000	-0.15

†$N_{Mal} = 13$, $N_{Hon} = 121$, $N_{M+H} = 134$

*$p < .01$

Table 3.15

Group Comparisons of Discrete Background Variables

Both Psychiatric and Neurocognitive Malingering vs. Honest (PNM vs. PNH)

		Both Psychiatric & Neurocognitive Malingering†		Both Psychiatric & Neurocognitive Honest†			
		N	%	N	%	χ^2	df
Demographics							
Race	White	13	100%	11	91.7%	1.161	2
	Af-Am	0	0.0%	1	7.4%		
	Other	0	0.0%	9	0.8%		
				1			
Gender	Male	7	53.8%	84	69.4%	1.307	1
	Female	6	46.2%	37	30.6%		
Marital Status	Single	2	15.4%	17	14.0%	0.825	4
	Married	8	61.5%	78	64.5%		
	Separated	0	0.0%	2	1.7%		
	Divorced	3	23.1%	19	15.7%		
	Widowed	0	0.0%	2	1.7%		
Claim Types							
Pain	Yes	3	23.1%	48	39.7%	1.078	1
Psychiatric	Yes	13	100.0%	88	72.7%	4.565	1
Brain Damage	Yes	2	15.4%	49	40.5%	3.140	1
Physical	Yes	4	30.8%	29	24.0%	0.274	1
Medical	Yes	1	7.7%	5	4.1%	0.338	1

†$N_{Mal} = 13$, $N_{Hon} = 121$, $N_{M+H} = 134$

*$p < .01$

Table 3.16

Group Comparisons of MMPI-2 Validity Scale Scores

Both Psychiatric and Neurocognitive Malingering vs. Honest (PNM vs. PNH)

	Both Psychiatric & Neurocognitive Malingering Group†			Both Psychiatric & Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
VRIN	13	49.31	9.214	118	51.19	8.321	-0.768	129	-0.22
TRIN	13	58.54	6.679	118	58.31	8.299	0.094	129	0.03
L	13	56.54	9.033	118	57.60	10.454	-0.352	129	-0.10
K	13	39.38	8.441	118	47.58	10.876	-2.629*	129	-0.77
S	13	38.31	8.625	118	47.64	10.385	-3.119*	129	-0.91
F	13	98.46	14.246	118	55.97	13.530	10.692*	129	3.13
Fb	13	113.54	10.952	118	60.09	17.242	10.914*	129	3.21
F(p)	13	65.08	11.398	118	50.21	10.189	4.935*	129	1.44
F – K	13	9.15	7.267	117	-6.94	11.038	5.125*	128	1.51

†$N_{Mal} = 13$, $N_{Hon} = 121$, $N_{M+H} = 134$

*$p < .01$

Table 3.17

Group Comparisons of MMPI-2 Clinical Scale Scores

Both Psychiatric and Neurocognitive Malingering vs. Honest (PNM vs. PNH)

	Both Psychiatric & Neurocognitive Malingering Group†			Both Psychiatric & Neurocognitive Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
1	13	91.54	7.230	118	75.75	12.694	6.801*	21.283	1.30
2	13	95.15	9.054	118	76.58	14.102	6.569*	19.131	1.36
3	13	92.38	9.314	118	75.97	16.311	5.495*	21.230	1.05
4	13	72.54	7.287	118	57.62	10.209	5.119*	129.000	1.50
5	13	51.46	8.771	118	46.97	8.605	1.784	129.000	0.52
6	13	85.00	20.789	118	57.12	14.517	6.273*	129.000	1.85
7	13	94.92	6.689	118	70.25	13.135	11.140*	23.917	1.97
8	13	101.92	6.751	118	64.24	14.840	16.259*	27.381	2.67
9	13	54.62	5.810	118	48.05	11.209	2.076	129.000	0.61
0	13	74.46	11.348	118	61.41	12.493	3.605*	129.000	1.05

†$N_{Mal} = 13$, $N_{Hon} = 121$, $N_{M+H} = 134$

*$p < .01$

Table 3.18

Group Comparisons of SIMS Total and Subscale Scores

Both Psychiatric and Neurocognitive Malingering vs. Honest (PNM vs. PNH)

	Both Psychiatric & Neurocognitive Malingering†			Both Psychiatric & Neurocognitive Honest†					
	N	M	SD	N	M	SD	t	df	*d*
SIMS									
Total	13	34.85	8.961	121	13.63	6.033	11.438*	132.000	3.37
N	13	8.54	3.099	121	3.12	2.399	7.521*	132.000	2.20
Af	13	9.08	1.382	121	5.84	2.327	7.386*	20.238	1.45
P	13	4.38	3.686	121	0.31	0.684	3.974*	12.089	4.23
LI	13	4.31	1.932	121	2.18	1.638	4.370*	132.000	1.28
Am	13	8.46	2.696	121	2.11	2.345	9.153*	132.000	2.67

†$N_{Mal} = 13$, $N_{Hon} = 121$, $N_{M+H} = 134$

*$p < .01$

Table 3.19

Group Comparisons of Continuous Background Variables

Any Malingering vs. Honest (AM vs. PNH)

	Any Malingering Group†			Both Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
Age	85	39.95	10.930	121	42.95	12.702	-1.764	204	-0.25
Education level	85	11.65	2.667	120	12.47	2.722	-2.142	203	-0.30

†$N_{Mal} = 85$, $N_{Hon} = 121$, $N_{M+H} = 206$

*$p < .01$

Table 3.20

Group Comparisons of Discrete Background Variables

Any Malingering vs. Honest (AM vs. PNH)

		Any Malingering†		Both Honest†			
		N	%	N	%	χ^2	df
Demographics							
Race	White	79	92.9%	111	91.7%	0.720	2
	Af-Am	6	7.1%	9	7.4%		
	Other	0	0.0%	1	0.8%		
Gender	Male	60	70.6%	84	69.4%	0.032	1
	Female	25	29.4%	37	30.6%		
Marital status	Single	12	14.1%	17	14.0%	1.378	4
	Married	52	61.2%	78	64.5%		
	Separated	1	1.2%	2	1.7%		
	Divorced	17	20.0%	19	15.7%		
	Widowed	3	3.5%	2	1.7%		
Claim Type							
Pain	Yes	36	42.4%	48	39.7%	0.129	1
Psychiatric	Yes	73	85.9%	88	72.7%	4.648	1
Brain damage	Yes	27	31.8%	49	40.5%	1.635	1
Physical	Yes	30	35.3%	29	24.0%	3.006	1
Medical	Yes	7	8.2%	5	4.1%	1.495	1

†$N_{Mal} = 85$, $N_{Hon} = 121$, $N_{M+H} = 206$

*$p < .01$

Table 3.21

Group Comparisons of MMPI-2 Validity Scale Scores

Any Malingering vs. Honest (AM vs. PNH)

	Any Malingering Group†			Both Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
VRIN	84	50.37	9.551	118	51.19	8.321	-0.639	163.133	-0.09
TRIN	85	58.28	7.179	118	58.31	8.299	-0.028	201.000	0.00
L	85	56.44	9.588	118	57.60	10.454	-0.812	201.000	-0.11
K	85	42.28	10.189	118	47.58	10.876	-3.518*	201.000	-0.50
S	85	42.80	11.132	118	47.64	10.385	-3.176*	201.000	-0.45
F	85	80.66	19.013	118	55.47	13.530	10.246*	142.816	1.59
Fb	85	90.56	23.537	118	60.09	17.242	10.136*	145.844	1.53
F(p)	85	58.27	13.018	118	50.21	50.21	4.754*	153.088	0.71
F – K	85	2.25	8.956	118	-6.94	-6.94	6.310*	200.000	0.90

†$N_{Mal} = 85$, $N_{Hon} = 121$, $N_{M+H} = 206$

**p* < .01

Table 3.22

Group Comparisons of MMPI-2 Clinical Scale Scores

Any Malingering vs. Honest (AM vs. PNH)

	Any Malingering Group†			Both Honest Group†					
	N	M	SD	N	M	SD	t	df	*d*
MMPI-2									
1	84	87.68	10.009	118	75.75	12.694	7.456*	197.864	1.03
2	84	92.70	11.611	118	76.58	14.102	8.886*	195.728	1.23
3	84	86.68	12.582	118	75.97	16.311	5.265*	198.670	0.73
4	84	75.10	80.412	118	57.62	10.209	2.337	200.000	0.44
5	84	49.55	7.872	118	46.97	8.605	2.176	200.000	0.31
6	84	75.29	20.091	118	57.12	14.517	7.076*	142.225	1.08
7	84	88.14	14.328	118	70.25	13.135	0.185*	200.000	1.31
8	84	87.80	17.970	118	64.24	14.840	10.180*	200.000	1.46
9	84	50.21	10.803	118	48.05	11.209	1.372	200.000	0.20
0	84	74.05	11.789	118	61.41	12.493	7.255*	200.000	1.04

† $N_{Mal} = 85$, $N_{Hon} = 121$, $N_{M+H} = 206$

**p* < .01

Table 3.23

Group Comparisons of SIMS Total and Subscale Scores

Any Malingering vs. Honest (AM vs. PNH)

	Any Malingering†			Both Honest†					
	N	M	SD	N	M	SD	t	df	*d*
SIMS									
Total	85	25.06	10.381	121	13.63	6.033	9.126*	123.717	1.46
N	85	6.38	3.512	121	3.12	2.399	7.429*	137.722	1.14
Af	85	8.05	2.187	121	5.84	2.327	6.859*	204.000	0.97
P	85	1.66	2.514	121	0.31	0.684	4.807*	92.772	0.94
LI	85	3.02	2.035	121	2.18	1.638	3.282*	204.000	0.47
Am	85	5.92	3.328	121	2.11	2.345	9.089*	140.797	1.39

†$N_{Mal} = 85$, $N_{Hon} = 121$, $N_{M+H} = 206$

*$p < .01$

Table 3.24

Summary of Test Score Difference Effect Sizes for Four Sets of Contrasts*

Test Score	PM vs. PH†	NM vs. NH†	PNM vs. PNH†	AM vs. PNH†
SIRS Total	6.79‡	0.99	7.34‡	2.02‡
MMPI-2 F	2.70	1.16	3.13	1.59
Fb	2.92	0.94	3.21	1.53
SIMS Total	2.55	1.13	3.37	1.46
TOMM Trial 2	-1.12	-3.03‡	-9.97‡	-2.04‡
VSVT Diff	-1.01	-4.92	-7.71	-3.41
Total	-1.10	-4.79‡	-8.56‡	-3.33‡
LMT	-1.25	-4.11‡	-10.71‡	-2.79‡

*Cohen's *d*

†PM = Psychiatric – Malingering, PH = Psychiatric – Honest, NM = Neurocognitive – Malingering, NH = Neurocognitive – Honest, PNM = Psychiatric and Neurocognitive – Malingering, PNH = Psychiatric and Neurocognitive – Honest, AM = Any – Malingering

‡These scores formed at least part of the criteria for group membership in this Comparison

Table 3.25

Summary of SIMS Score Difference Effect Sizes for Four Sets of Contrasts*

Test Score	PNM vs. PNH†	PM vs. PH†	AM vs. PNH†	NM vs. NH†
SIMS Total	3.37	2.55	1.46	1.13
N	2.20	1.85	1.14	0.80
Af	1.45	1.42	0.97	0.61
P	4.23	3.27	0.94	0.82
LI	1.28	0.94	0.47	0.41
Am	2.67	1.77	1.39	1.23

Note: Contrasts are presented from largest effect sizes on far left to smallest on far right.

*Cohen's *d*

†PM = Psychiatric – Malingering, PH = Psychiatric – Honest, NM = Neurocognitive – Malingering, NH = Neurocognitive – Honest, PNM = Psychiatric and Neurocognitive – Malingering, PNH = Psychiatric and Neurocognitive – Honest, AM = Any – Malingering

Table 3.26

Conditional Stepwise Regressions of Criterion Measures Onto SIMS Total and MMPI-2 Infrequency Scales*

Regress	Step	Onto	R	R^2	ΔR^2	ΔF	Sig
SIRS Total	1	SIMS Total	0.707	0.499	0.499	299.372	0.000
	2	MMPI Fb	0.780	0.608	0.109	83.166	0.000
	3	MMPI F	0.785	0.616	0.008	6.080	0.014
	1	MMPI Fb	0.726	0.527	0.527	334.333	0.000
	2	MMPI F	0.744	0.554	0.027	17.809	0.000
	3	SIMS Total	0.785	0.616	0.063	48.594	0.000
TOMM T2	1	SIMS Total	0.402	0.161	0.161	57.686	0.000
	1	MMPI F	0.330	0.109	0.109	36.705	0.000
	2	SIMS Total	0.410	0.168	0.059	21.268	0.000
VSVT Diff	1	SIMS Total	0.446	0.199	0.199	74.516	0.000
	2	MMPI F	0.486	0.236	0.037	14.435	0.000
	1	MMPI F	0.441	0.194	0.194	72.412	0.000
	2	SIMS Total	0.486	0.236	0.041	16.206	0.000
VSVT Total	1	SIMS Total	0.453	0.205	0.205	77.271	0.000
	2	MMPI F	0.488	0.238	0.033	12.951	0.000
	1	MMPI F	0.438	0.191	0.191	71.047	0.000
	2	SIMS Total	0.488	0.238	0.046	18.183	0.000
LMT	1	SIMS Total	0.443	0.197	0.197	73.434	0.000
	2	MMPI F	0.461	0.213	0.016	6.150	0.014
	1	MMPI F	0.391	0.153	0.153	54.135	0.000
	2	SIMS Total	0.461	0.213	0.060	22.779	0.000

*Probability of F to enter ≤ .050; probability of F to remove ≥ .100

Table 3.27

Classification Accuracy Statistics Using Cut Score of SIMS Total > 14

	Comparison				Edens et al.	Smith & Burger
	Psych	NC	Both	Any	(1999)	(1997)
BR	7.5%	24.4%	4.2%	27.6%	50.0%	85.7%
Sensitivity	95.7%	80.0%	100.0%	81.2%	96.0%	96.0%
Specificity	54.1%	47.2%	57.0%	57.0%	91.0%	88.0%
HR	56.8%	56.9%	61.2%	67.0%	96.0%	95.0%
PPP	21.8%	39.0%	20.0%	57.0%	91.4%	98.0%
NPP	98.9%	84.8%	100.0%	81.2%	95.8%	78.6%
IPPP	14.3%	14.6%	15.8%	29.4%	41.4%	12.2%
INPP	6.4%	9.2%	4.2%	8.8%	45.8%	64.3%

Table 3.28

Classification Accuracy Statistics Using Cut Score of SIMS Total > 16

	Comparison				Lewis et al.	Rogers et al.	Smith
	Psych	NC	Both	Any	(2002)	(1996)	(1992)
BR	7.5%	24.4%	4.2%	27.6%	31.4%	50.0%	87.5%
Sensitivity	95.7%	74.7%	100.0%	76.5%	100.0%	●*	●*
Specificity	66.9%	60.1%	70.3%	70.3%	61.0%	●*	●*
HR	68.7%	64.4%	73.1%	72.8%	73.0%	●*	95.0%
PPP	27.9%	44.1%	26.5%	64.4%	54.0%	87.0%	●*
NPP	99.1%	84.9%	100.0%	81.0%	100.0%	62.0%	●*
IPPP	20.4%	19.7%	22.3%	36.8%	22.6%	37.0%	●*
INPP	6.6%	9.3%	4.2%	8.5%	31.4%	12.0%	●*

*● = Figures not published

Table 3.29

Cut Scores Yielding Maximum Sensitivity and NPP in NC and Any Comparisons

Comparison: Cut Score:	NC > 4	Any > 4
BR	24.4%	27.6%
Sensitivity	100.0%	100.0%
Specificity	51.0%	58.0%
HR	33.2%	44.7%
PPP	39.6%	47.6%
NPP	100.0%	100.0%
IPPP	15.3%	20.0%
INPP	24.4%	27.6%

Chapter Four

Discussion

Background

Malingering presents a significant problem for society by facilitating fraud and the avoidance of criminal responsibility. Any psychological examination conducted in a compensation-seeking context should include an assessment of the likelihood of malingering (Rogers, 1997). Historically, scales embedded in instruments already commonly included in forensic assessment batteries were validated and employed in this effort; but difficulties in establishing reliable cut scores that were useful across samples (Lewis et al., 2002; Inman & Berry, 2000) led researchers to develop dedicated malingering measures.

The SIRS (Rogers, et al., 1992) is considered the gold standard for the detection of psychiatric symptom malingering (Lewis et al., 2002); however, it requires 30 to 60 minutes of one-on-one professional administration time, and has been designed to minimize false positives. Therefore, a brief screening instrument, designed to minimize false negatives, could save considerable time and expense by eliminating clearly honest examinees from further testing, at the same time increasing the base rate of feigning in those who proceed to subsequent evaluation with the SIRS. The already impressive PPP of the SIRS would, consequently, increase, and the overall accuracy of the two-stage classification system would be substantially higher than that of the SIRS alone.

The SIMS (Smith & Burger, 1997) was designed to fulfill the role of a first-stage screening measure of malingering of various types of symptomatology. Initial validation studies have demonstrated preliminary evidence of its utility in the detection of psychiatric symptom malingering. Four simulation studies (Smith, 1992; Rogers et al., 1996; Smith & Burger, 1997; Edens et al., 1999), one known-group analysis (Heinze & Purisch, 2001), and one known-groups comparison (Lewis et al., 2002) have been conducted, showing promise for its role as a high-NPP test of malingering. A consistent cut score is not clearly suggested by the existing literature. In addition, only one simulation study (Rogers et al., 1996) included patient or offender participants; all studies that did include offenders (Rogers et al., 1996; Heinze & Purisch, 1999; Lewis et al., 2002) involved small, all-male samples; only the report on the known-groups comparison by Lewis et al. (2002) provided all classification accuracy statistics; and many authors failed to indicate how reported cut scores were selected. No samples were derived from a civil forensic setting, and none directly tested the SIMS' ability to detect neurocognitive symptom feigning. Effect sizes, reported only for the known-groups comparison, ranged from 1.1 to 3.0 (Lewis et al., 2002).

In the present study, 308 plaintiffs in workers' compensation, personal injury, and fitness for duty cases were assessed with the SIRS; three well-validated tests of neurocognitive effort (the TOMM, developed by Tombaugh, 1996; VSVT, created by Slick et al., 1997; and LMT, designed by Inman et al., 1998); the SIMS; and a background questionnaire. All but six of these participants also completed the MMPI-2 (Butcher et al., 1989). Four pairs of comparison groups were created: a group determined to be malingering psychiatric symptoms (according to the SIRS) and a corresponding group determined to be honestly responding with respect to psychiatric symptoms (according to the SIRS); a group classified as malingering neurocognitive deficits (according to at least two of three tests of neurocognitive effort) and a corresponding group classified as honest responders with respect to neurocognitive deficits (according to all three tests of neurocognitive effort); a group considered to be malingering both psychiatric and neurocognitive symptoms (by meeting all criteria which define the first two malingering groups) and a corresponding group of individuals considered to be honestly responding with respect to both psychiatric and neurocognitive symptoms (by meeting all criteria which define the first two "honest" groups); and, finally, a group categorized as malingering any symptomatology – psychiatric, neurocognitive, or both (according to the SIRS and/or two of the three tests of neurocognitive effort), which was also compared to the group categorized as honest with respect to both psychiatric and neurocognitive symptom reporting (again, by meeting all criteria which define the first two "honest" groups). Each pair of malingering and honest groups was compared with respect to demographics, types of claims made, and test scores on all measures listed. The ability of the SIMS to predict scores produced by established malingering measures was assessed via AUC computations and linear regression analyses. And, ultimately, in order to evaluate the SIMS' utility in detecting individual malingerers, classification accuracy statistics were generated for each pair of malingering and honest groups. Calculations indicate that power of the four subsamples in the current study to detect effect sizes of 1.0 (which is smaller than any effect sizes previously reported) ranges from 79.02% to 100.00%; power to detect effect sizes of 2.0 or 3.0 is 100.00% across all subsamples.

Differences Between Malingerers and Honest Responders

Malingering, in this sample, is not related to demographic variables or to the types of claims registered by the plaintiff. Demographic variables included gender, race, age, educational level, and marital status; claim types included psychiatric, brain damage, pain, medical, and physical, and many cases involved multiple claim types.

Corresponding malingering and honest groups differed substantially, however, on most test scores. Of particular interest, SIRS scales and Total, TOMM Trial 2, VSVT Difficult and Total, LMT (each of which, it must be noted, served as or was confounded with a criterion measure in three of the four comparisons), and SIMS scale and Total effect sizes were generally *quite* large, and, most notably, only three of twenty-four SIMS effect sizes were *not* large (less than .80; Cohen, 1977); sixteen (two-thirds) of them were larger than 1.00. SIMS effect sizes ranged from a low of .41 to a high of 4.23. These data lend substantial support to the supposition that the SIMS is measuring symptom exaggeration and/or feigning.

Evidence for Broadband Symptom Exaggeration

Interestingly, psychiatric symptom malingerers performed more poorly on all neurocognitive tests of effort than honest psychiatric responders. Correspondingly, neurocognitive symptom feigners endorsed more psychiatric symptoms than honest neurocognitive test responders. In this sample, then, malingerers as a group seemed to exaggerate symptoms “across the board”. This suggests that psychiatric malingerers may attempt to appear to experience mild neurocognitive impairment, although usually without triggering detection by neurocognitive feigning tests, while, similarly, malingerers of neurocognitive deficit elevate scales by endorsing psychiatric symptoms, but at a sub-clinical level such that most are not detected as psychiatric symptom malingerers. If replicated in other samples, this provides a hint of potentially useful information about approaches employed by malingerers in compensation-seeking settings that could influence malingering-detection strategies in the future.

Prediction of Criterion Scores by SIMS Total

At a group level, the SIMS Total score exhibited an ability to predict scores on previously-validated tests of malingering, and demonstrated incremental utility over the MMPI-2 Infrequency scales in this respect.

Validation of Previously Recommended Cut Scores

Previously recommended cut scores for the SIMS Total were predictive of malingering in this sample according to all four definitions of the construct implemented in this study. Sensitivity values using a cut score of > 14 to screen for *psychiatric* malingering were very high, and comparable to rates found in studies by Edens et al. (1999) and Smith & Burger (1997). NPP values were very high in these comparisons as well, comparable to results found by Edens et al. (1999) and much higher than those reported by Smith & Burger (1997). Detection of neurocognitive malingering produced appreciably lower sensitivity rates and NPPs.

Using a cut score > 16, sensitivity values for the detection of psychiatric malingering were very high, and comparable to Lewis et al.'s (2002) finding. NPP values were also very high in these comparisons, as were results found by Lewis et al. (2002), and were much higher than NPP figures reported by Rogers et al. (1996). Detection of neurocognitive malingering, however, produced substantially lower sensitivity rates and NPPs.

Support offered by the present study for both the > 14 and > 16 cut scores for the detection of psychiatric symptom malingering is encouraging. Although neither of these cut scores provided for *maximal* sensitivity or NPP for both PM-PH and PNM-PNH comparisons, all sensitivities and NPPs were quite high (ranging from .957 to 1.000 and .989 to 1.000 for sensitivity and NPP, respectively) in groups composed of psychiatric feigners. A cut score of > 16 yielded higher specificities and PPPs, and therefore its performance was superior to a cut of > 14 in this sample.

However, these cut scores performed less adequately in the detection of neurocognitive malingering: sensitivity values ranged from only .747 to .812 and, at the base rates in this sample, NPPs ranged from only .810 to .849. Consequently, the cut score that produced maximum sensitivity and NPP when screening for neurocognitive malingering in this sample was determined. This optimal cut score, > 4, is far below both of the previously recommended cut scores for psychiatric symptom malingering, and is of only minimal utility, as it screens out few honest responders. Reasons for the divergence between this cut score and scores that are effective at screening for psychiatric malingering are not clear. One possible factor could be method variance – all of the tests of neurocognitive effort are computer-administered, multiple-alternative, forced-choice tests of recognition memory; the SIRS is a structured interview comprised of questions about symptom endorsement, which are far more similar to SIMS items. The SIMS, a paper-and-pencil, true/false, self-report questionnaire, may be more effective at predicting results of one type of instrument as compared with the other.

The usefulness of the SIMS in the detection of the malingering of neurocognitive impairment, then, has yet to be demonstrated. More promising results were expected in light of the inclusion of scales in the SIMS that were specifically designed to tap malingering of Neurologic impairment (N), Amnestic symptoms (Am), and Low Intelligence (LI). The SIMS cannot be recommended for use in screening for neurocognitive symptom feigning until a cut score that consistently rules out a significant number of non-malingerers, and not true malingerers, is established across samples. It may be that different cut scores are applicable for different settings or applications; for example, it is possible that a lower cut score is useful when neurocognitive deficit is feigned, while a higher cut score performs reliably with individuals feigning psychiatric problems. Only further studies will provide the information needed to evaluate such possibilities.

Is the SIMS Ready for Forensic Use?

This study shows initial support for the utility of the SIMS in a civil forensic sample of men and women – yet, again, more research is needed in similar samples. Further studies could also help researchers evaluate the SIMS' potential in screening for neurocognitive symptom malingering, including its amenability to the establishment of an appropriate cut score for this purpose.

Although support for the SIMS is demonstrated by only a handful of studies, its use in some forensic settings might appear reasonable. Because the SIMS is designed for use only as a screen, it could be administered with some confidence that it could increase the utility of a second-stage assessment without substantial risk of failing to "capture" a true malingerer. This would, however, presume selection of a cut score that has consistently, across studies in samples similar to the intended setting, demonstrated high sensitivity and NPP. Until now, no more than two studies have even *reported* sensitivity *and/or* NPP values for the same cut score. It is on this basis that caution is recommended in employing the SIMS in forensic assessments at this time.

Study Limitations

As the implications of these results are evaluated, limitations of the present study must qualify any conclusions reached. Importantly, the nature of the known-groups design prevents us from definitively attributing differences between designated malingerers and honest responders to their malingering statuses (Rogers, 1997). Although potentially confounding demographic and claim type variables did not differ between any pair of malingering and honest groups, there may be other differences between the groups that were not evaluated that account for some or all of the effects detected. Notably, the success of the SIMS in several previous simulation studies complements and adds credibility to the results of this and previous known-groups research.

As with any known-groups comparison, to the extent that the validity of the criterion measures used is less than perfect, confidence that the malingering and honest designations were accurately assigned is less than 100%. As a result, the validity of the criterion measures (the SIRS, TOMM, VSVT, and TOMM) places a ceiling on the validity that can be established for the SIMS. Although all of the criterion measures are supported by substantial research demonstrating their reliability, validity, and utility in discriminating malingerers from honest responders, none of these instruments is a perfect indicator.

The sample, while quite large, included approximately twice as many men as women, and was overwhelmingly Caucasian, with virtually no racial variation among the non-Caucasian participants. Although geographic information was not available, it is highly likely that most of the participants resided in Kentucky, with the rest residing in nearby states. These restrictions in the demographic variability of the sample may yield different results than would a more diverse sample.

Suggestions for Future Research

The SIMS shows considerable potential in its intended role as a brief screen for malingering; however, results published so far are not sufficient to lend conclusive support to the practice of relying on the use of the SIMS in forensic evaluations which contribute so substantially to consequential outcomes. Future research needed in order to demonstrate the utility of the SIMS in forensic assessments includes the following.

Known-Groups Comparisons. The present study is only the second known-groups comparison supporting the utility of the SIMS. More such studies (with high generalizability to malingering in real-world settings) are essential, to complement the more numerous simulation studies that have been conducted.

Civil Forensic Samples. This is the first study evaluating the performance of the SIMS in a civil forensic sample. More research in similar settings could help to validate use of the SIMS with this population, and will also provide more information regarding typical base rates in civil forensic settings. It is also important that both male and female participants be included.

Investigation of the SIMS' Utility in Screening for Neurocognitive Symptom Feigning. The SIMS contains three scales designed to be sensitive to the malingering of neurocognitive symptomatology, yet its ability to screen for this type of malingering has not been a focus of previous research, and has not been investigated in a known-groups comparison until now. Further assessments of the SIMS' performance in the detection of the feigning of neurocognitive impairment are recommended.

More Criminal Forensic Samples. Although three studies have included offender participants, one (Rogers et al., 1996) was a within-subjects simulation study conducted with a sample of male adolescent offenders; and another (Heinze & Purisch, 2001) only assessed the sensitivities of several instruments in a single group of male offender malingerers identified according to questionable criteria. Additional studies in criminal forensic samples, including both men *and* women; are sorely needed in order to strengthen evidence for the utility of the SIMS in such settings.

Utility of SIMS Subscales. Few data have been published documenting the sensitivity of individual subscales to various types of feigning that include complete classification accuracy statistics for the subscales or that evaluate the utility of subscales used alone or in various combinations. Although in the current investigation data relevant to such an investigation were collected, these questions were not explored here. Examinations of the performance of each SIMS subscale would allow for the evaluation of each scale's usefulness in tapping malingering of the type of symptomatology it was designed to assess, and might reveal screening strategies superior to the use of the Total score in some settings or samples.

Investigation of SIMS' Performance in Screening for Pain Malingering. Detection of pain malingering is a relatively recent area of investigation (Rogers, 1997). Assessment of the utility of the SIMS, as well as other instruments, in screening for the malingering of pain is suggested.

Diverse Demographic Samples. The current study utilized a sample that was quite restricted racially and geographically; such lack of demographic variability is a common issue, but raises questions about the generalizability of results to individuals and samples who do not resemble study participants. Evaluations of the performance of the SIMS with diverse demographic samples are encouraged.

Validation of Reliable Cut Scores. Most critically, work must be undertaken toward the establishment of a reliable cut score, or of reliable cut score*s* for specific applications. Additional efforts to validate the previously proposed cut scores of > 14 and, especially, > 16 should be attempted so that adequate evidence for a cut score that is consistently useful across *or at least within certain* sample types can be amassed.

Conclusion

This study adds to existing support for the SIMS' potential. The SIMS shows continued promise as a brief screen for malingering, but requires consistent and well-supported cut scores before it can be relied upon to assist in the high-stakes decision-making performed as a matter of course in forensic contexts.

References

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th Ed., Text Revision). Washington, D.C.: Author.

Arbisi, P. A. & Ben-Porath, Y. S. (1998). The ability of Minnesota Multiphasic Personality Inventory – 2 validity scales to detect fake-bad responses in psychiatric inpatients. *Psychological Assessment, 10,* 221-228.

Berry, D. T. R., Baer, R. A., & Harris, M. J. (1991). Detection of malingering on the MMPI: A meta-analytic review. *Clinical Psychology Review, 11,* 585-598.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences* (Revised Ed.). New York: Academic Press.

Derogatis, L.R. (1977). *SCL-90-R Administration, Scoring, and Procedures Manual*. Minneapolis: NCS.

Edens, J. F., Otto, R. K., & Dwyer, T. (1999). Utility of the Structured Inventory of Malingered Symptomatology in identifying persons motivated to malinger psychopathology. *Journal of the American Academy of Psychiatry and the Law, 27,* 387-396.

Heinze, M. C. & Purisch, A. D. (2001). Beneath the mask: Use of psychological tests to detect and subtype malingering in criminal defendants. *Journal of Forensic Psychology Practice, 1,* 23-52.

Inman, T. H. & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology, 17,* 1-23.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10,* 128-139.

Lewis, J. L., Simcox, A. M., & Berry, D. T. R. (2002). Screening for Feigned Psychiatric Symptoms in a Forensic Sample by Using the MMPI-2 and the Structured Inventory of Malingered Symptomatology. *Psychological Assessment, 14,* 170-176.

Meehl, P. E. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52,* 194-216.

Rogers, R. (Ed.). (1997). *Clinical Assessment of Malingering and Deception* (2nd Ed.). New York: Guilford Press.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *SIRS - Structured Interview of Reported Symptoms: A professional manual.* Odessa, FL: Psychological Assessment Resources Inc.

Rogers, R., Hinds, J. D., & Sewell, K. W. (1996). Feigning psychopathology among adolescent offenders: Validation of the SIRS, MMPI-A, and SIMS. *Journal of Personality Assessment, 67,* 244-257.

Rogers, R., Sewell, K. W., & Salekin, R. T. (1994). A meta-analysis of malingering on the MMPI-2. *Assessment, 1,* 227-237.

Slick, D., Hopp, G., Strauss, E., & Thompson, G. B. (1997). *Victoria Symptom Validity Test.* Odessa, FL: Psychological Assessment Resources Inc.

Smith, G. P. (1992). Detection of Malingering: A validation study of the SLAM test (Doctoral dissertation, University of Missouri - St. Louis, 1992). *Dissertation Abstracts International, 53,* 3795B.

Smith, G. P. & Burger, G. K. (1997). Detection of Malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law, 25,* 183-189.

Thompson, G. B. (2002). The Victoria Symptom Validity Test: An enhanced test of symptom validity. *Journal of Forensic Neuropsychology, 2,* 3-4, 43-67.

Tombaugh, T. N. (1996). *Test of Memory Malingering: TOMM.* North Towanda, NY: Multi-Health Systems.

Tombaugh, T. N. (2002). The Test of Memory Malingering (TOMM) in forensic psychology. *Journal of Forensic Neuropsychology, 2,* 3-4, 69-96.

Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E., & Orey, S. A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology, 19,* 37-48.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – Third Edition.* San Antonio, TX: The Psychological Corporation.

Vita

Yvonne Renee Alwes

Personal Information

Date of Birth: May 6, 1959

Place of Birth: Ephrata, Pennsylvania

Education

Ph.D. Program, Clinical Psychology. August 2003 – Present.
University of Kentucky, Lexington, Kentucky.

Bachelor of Arts, Psychology. May 1981.
Transylvania University, Lexington, Kentucky.

Scholastic Honors

National Merit Scholarship
Recipient

Omicron Delta Kappa, National Leadership Honorary
Member and Vice President of University Chapter

Experience

Psychological Assessment.
University of Kentucky. August 2003 – Present.

Psychotherapy.
Jesse G. Harris Psychological Services Center, University of Kentucky.
August 2004 – Present.

Professional Memberships

Association for Psychological Science
(formerly American Psychological Society)

Society for a Science of Clinical Psychology
(American Psychological Association — Division 12 — Section III)

American Psychology - Law Society
(American Psychological Association — Division 41)

Kentucky Psychological Association

Yvonne Renee Alwes